Molecular patterns in acute pancreatitis reflect generalisable endotypes of the host response to systemic injury in humans

Molecular patterns in acute pancreatitis reflect generalisable endotypes of the host response to systemic injury in humans.

Authors

Lucile P. A. Neyton[1,2#], Xiaozhong Zheng[1], Christos Skouras[3], Andrea Doeschl-Wilson[2], Michael U. Gutmann[4], Iain Uings[5], Francesco V. Rao[6], Armel Nicolas[6], Craig Marshall[7], Lisa-Marie Wilson[7], J. Kenneth Baillie[2*#] and Damian J. Mole[1,3*#].

Affiliations

[1]Medical Research Council Centre for Inflammation Research, Queen's Medical Research Institute, The University of Edinburgh, Edinburgh, UK

[2]Division of Genetics and Genomics, The Roslin Institute, The University of Edinburgh, Edinburgh, UK

[3]Clinical Surgery, School of Clinical Sciences and Community Health, The University of Edinburgh, Edinburgh, UK

[4]School of Informatics, The University of Edinburgh, Edinburgh, UK

[5]GSK Discovery Partnerships with Academia, Exploratory Discovery, Future Pipeline Discovery, Medicines Research Centre, Gunnels Wood Road, Stevenage, UK.

[6]DC Biosciences Limited, James Lindsay Place, Dundee Technopole, Dundee, UK

[7]Department of Laboratory Medicine, NHS Lothian, Edinburgh, UK.

* indicates joint senior authors

[#] indicates corresponding authors

Corresponding authors

j.k.baillie@ed.ac.uk; damian.mole@ed.ac.uk

Keywords

Acute pancreatitis, endotypes, critical illness trajectory, multi-omics, time series, cluster analysis

Abstract

Objective: Acute Pancreatitis (AP) is sudden onset pancreas inflammation that causes systemic injury with a wide and markedly heterogeneous range of clinical consequences. Here, we hypothesised that this observed clinical diversity corresponds to diversity in molecular subtypes that can be identified in clinical and multi-omics data.

Summary Background data: Observational cohort study. n=57 for the discovery cohort (clinical, transcriptomics, proteomics and metabolomics data) and n=312 for the validation cohort (clinical and metabolomics data).

Methods: We integrated co-incident transcriptomics, proteomics, and metabolomics data at serial time points between admission to hospital and up to 48 hours after recruitment from a cohort of patients presenting with acute pancreatitis. We systematically evaluated four different metrics for patient similarity using unbiased mathematical, biological and clinical measures of internal and external validity.

We next compared the AP molecular endotypes with previous descriptions of endotypes in a critically ill population with acute respiratory distress syndrome (ARDS).

Results: Our results identify four distinct and stable AP molecular endotypes. We validated our findings in a second independent cohort of patients with AP.

We observed that two endotypes in AP recapitulate disease endotypes previously reported in ARDS.

Conclusions: Our results show that molecular endotypes exist in AP and reflect biological patterns that are also present in ARDS, suggesting that generalisable patterns exist in diverse presentations of critical illness.

Introduction

AP is defined as acute inflammation of the pancreas(1) and has a worldwide incidence of 34 per 100,000 person-years(2). It is the commonest gastrointestinal cause for emergency hospital admission(3). Inflammatory damage to pancreatic acinar cells initiates an inflammatory cascade mediated by damage-associated molecular patterns, alarmins, inflammatory cytokines, metabolites and other soluble and cellular mediators of inflammation that propagate inflammation locally in the pancreas, and cause extrapancreatic organ dysfunction in the lungs, kidney, liver and other body systems, together resulting in multiple organ dysfunction syndrome (MODS)(4, 5). MODS occurs in 1 in 4 individuals who develop AP and is accompanied by deregulation of cardiovascular, autonomic nervous and immune system homeostasis(6), leading to death in one fifth of those with AP-MODS(7). Despite this currently accepted unifying disease model, clinical patterns of AP are markedly heterogeneous and severity is not directly proportional to the amount of pancreas damage on radiological imaging(8, 9). AP is caused by a wide range of precipitants, including choledocholithiasis, excess ingestion of alcohol, trauma, pancreatic manipulation at endoscopy, viral infections, certain venoms and specific prescription medicines(6). Currently, the clinicopathological paradigm in AP is convergent: diverse etiologies converge onto acinar cell damage, and the resulting systemic inflammatory response is stratified as mild, moderate or severe. In common with other types of systemic injury(10-16), we hypothesized the existence of molecular subtypes in AP, designated as endotypes(17). We predict that detailed knowledge of those endotypes will have clinical and therapeutic relevance(18). As new medicines for AP emerge, the existence of endotypes will become critical to directing treatment choices and understanding individualised responses to therapy.

The methodology to tackle this problem, particularly using time-series multi-omics data, is not settled. We designed a systematic, data-driven evaluation of four different metrics to quantify dissimilarity between patients and cluster them accordingly.

Materials and Methods

Data acquisition

We used samples and data from the IMOFAP (Inflammation, Metabolism and Organ Failure in AP) cohort(19) (n=79 recruited participants) as our discovery set and samples and data from n=312 patients from the KAPVAL (Kynurenine pathway in AP, VALidation) cohort, as the validation set (full details of the sampling and cohorts in **Supplementary Materials and Methods**).

AP diagnosis was confirmed for 57 out of 79 recruited patients in IMOFAP according to the revised Atlanta criteria(20). Integrated multi-omics analysis excluded 3 patients who had a prolonged interval (>200 hours) after symptom onset and therefore were late in their disease trajectory to avoid bias (**Figure 1a**). The median time interval between symptom onset and recruitment was 21.3 hours (Q1-Q3 13.5-54.4 in hours). Multiple timed samples of peripheral venous blood were taken (**Figure 1b**). When analysing single time point data from time point 0, we selected 40 patients (22 mild, 14 moderate and 4 severe AP cases) based on whether they had a complete multi-omics dataset. When comparing time-series, we required at least two complete time points, therefore selecting 34 patients (16 mild, 13 moderate and 5 severe AP cases). The KAPVAL validation cohort included 312 participants (274 in ward care level, 7 admitted to a high-dependency unit, and 31 to an intensive care unit) with confirmed AP and clinical data and serum metabolome at the time of initial presentation to hospital.

Data pre-processing is summarised in **Supplementary Figure 1** and **Supplementary Materials and Methods** for transcriptomics, proteomics and metabolomics data.

Data Analysis

We considered and evaluated four data analysis strategies. We tested single time point-based strategy using Euclidean distances and three time-series-based methods: 1) Area Under the Curve and PCA, 2) PCA with Trajectory(21) and 3) Dynamic Time Warping(22) (described in more detail in **Supplementary Materials and Methods**). For all methods, we used pre-processed Z-scores for the concatenated dataset (metabolomics, proteomics and transcriptomics) to take full advantage of the breadth of available measurements.

Single time point Euclidean distances

The first considered strategy consisted of computing Euclidean between all pairs of patients at multiple time points. This was performed for time point 0, 24 and 48 hours individually and obtained measures were used to quantify dissimilarity between individuals.

AUC and PCA

For each variable and individual we computed area under the curve (AUC) values for the corresponding time series using the trapezoidal rule (**Supplementary Materials and Methods)** and projected this new dataset onto Principal Component (PCA) space where, using the first two components, we computed Euclidean distances between individuals (**Supplementary Materials and Methods**).

Trajectory through PCA space

Trajectories of patients through PCA space were used to cluster patients(21), projecting all time points onto a two-dimensional PCA space, using the first two components. To map the trajectory evolution through PCA space we considered the direction between pairwise time points for each individual using integer values between 1 and 4 and the Hamming distance(21) to evaluate the dissimilarities between individuals.

Dynamic time warping

Dynamic time-warping (DTW) (22) is an algorithm that aligns and compare time series, by warping the time axis, finding the best alignment and the distance between two transformed curves. We compared patients using DTW, considering all pre-processed variables and estimated distances between pairs of individuals.

Clustering

We sought to discover AP endotypes within these data, specifically seeking disease trajectory groupings. To achieve this, we applied unsupervised clustering to patient-to-patient distance matrices(23), and devised a systematic, data-driven evaluation of the results from the different methods.

Using the dissimilarity matrices previously obtained we clustered them using hierarchical clustering and Ward's method (**Supplementary Materials and Methods)**.

The number of clusters was chosen according to stability, computed by generating 100 bootstrapped sets, clustering them and assessing the overlap with the original solution. Robustness of these clusters and the potential to be generalisable to other cohorts was estimated (**Supplementary Materials and Methods**).

Validation

To perform the validation, we selected time point 0 values to mimic a realistic situation in which individuals would be allocated to a subgroup as early as possible in their disease trajectory.

As bootstrapping with cluster comparisons produced a measure of stability, we used it to filter solutions based on a Jaccard index (JI) threshold of 0.75(24).

Compound set enrichment analyses were carried out to assess biological plausibility. Our working model was: each clustering method is intended to detect real biological similarities among a group of patients, rather than chance groupings due to random noise. We predicted that a method that detects real similarities between patients is more likely to recapitulate known biological groupings that drive the clustering. KEGG(25) data was used to perform the enrichment and test subsets of compounds. Generalised linear models were generated for each compound to test whether they identified groups. Results for all compounds of a set were summarised and significant items (using a p-value threshold of .05) used to quantify biological relevance (**Supplementary Materials and Methods**)

Enrichment

Partial Least Square Discriminant Analysis (PLS-DA) was used to highlight biomarkers within each subgroup, weighted and used to compute Variable Importance in Projection (VIP) scores. Applying a threshold of 1, top variables were retained for the enrichment using Reactome data and Fisher's exact test (**Supplementary Materials and Methods**).

Validation in an independent dataset

To validate the subgroups, we used the KAPVAL cohort clinical data and metabolomes, excluding drug metabolites, normalised as done for IMOFAP. To detect data structural similarity,

we classified KAPVAL samples using PLS-DA models with 25 or less metabolite predictors (**Supplementary Materials and Methods**). Signals between IMOFAP and samples from the KAPVAL cohort allocated to these same groups were compared using metabolites not included in the models (369 metabolites). Additionally, we computed in-group proportions (IGP) and associated FDR-corrected p-values(28). Briefly, the IGP of an endotype is the proportion of samples allocated to that endotype having their nearest neighbour (determined using Pearson's correlation coefficients) allocated to that same endotype.

Comparison with another disease

We compared our endotypes with ARDS endotypes described in another study(11). From ARDS endotypes we extracted variables rankings and compared it to rankings in each one of our endotypes. For each one of the two ARDS cohorts (ALVEOLI and ARMA), we compared the lists of ranked variables to each one of our endotypes using Spearman's correlation coefficients and FDR-corrected p-values (**Supplementary Materials and Methods).**

Comparison with an independent tool

MOFA*(29)* highlights variables explaining variation between samples, using factor analysis. Model factors explaining less than 1% of the variance in all omics were dropped. We clustered individuals into a 4-cluster solution using two latent factors and compared MOFA to previously generated outputs.

Results

An overview of time-series analysis-based results is presented in **Figure 1c**.

Clustering

The biology of inflammation in AP is complex and, when the entire AP cohort is taken as a whole, is seemingly heterogeneous. In order to uncover otherwise hidden groupings within the overall patient group with AP, seeking similarity between individuals who are part of an apparently dissimilar larger group, we integrated transcriptomics, proteomics and metabolomics data to have a comprehensive view of the biological processes across these different systems layers(30). Because the majority of patients who develop AP-MODS are admitted to critical care within 48 hours after presentation to hospital(7), we included time points between 0 and 48 hours after presentation (**Figure 2b**). We took an unsupervised approach to the statistical analysis, i.e. without preconceived notions of expected dominance of certain molecular mechanisms, and we initially included all variables that were available to us to avoid bias due to previous findings or hypotheses. After data pre-processing (**Supplementary Figure 1**), we created a combined dataset consisting of 651 metabolites, 371 proteins and 19766 genes that was used as input. The area under the curve combined with PCA, trajectory through PCA space and dynamic time warping methods required more than a time point per sample and thus 20 were discarded, resulting in 34 samples being analysed (5 severe, 13 moderate and 16 mild AP). Demographics are presented in **Supplementary Table 1**.

To highlight subgroups of patients of potential interest, once dissimilarity matrices were obtained, hierarchical clustering with the agglomerative Ward's method(31-33) was applied (**Figure 1c**).

Internal validity

Next, to ensure that the groupings that we had discovered were robust and stable, in other words would not be drastically altered if one or two individuals or variables changed, we used the Jaccard Index (JI). The stability criterion (JI>0.75) was met by the AUC in PCA space method (JI=0.79) and state-space trajectories in PCA space (JI=0.76), demonstrating internally robust groups (**Figure 1c and Supplementary Figure 2, Supplementary Tables 2 and 3**) whose structure would be expected to be recovered using a different input set.

Biological validity, or the number of known biological processes that were discretely identified by the clusterings, was strongest for the AUC in PCA space clustering method over alternatives, based on number of FANTOM5 and KEGG gene-based and KEGG metabolite-based hits (**Figure 1c, Supplementary Table 4)**. We therefore took forward AUC in PCA space as the solution for external validation.

Endotype characterisation

In order to explore the biological relevance of the discovered endotypes, and, also if maximum clinical utility is to be obtained from endotype assignation in AP, the endotype should be identifiable as close as possible to the time the affected individual seeks medical help.

We identified each one of our four endotypes using capital letters, A (n=13), B (n=10), C (n=5) and D (n=6). Each identified top variable (**Figure 2a),** here either gene or metabolite, was cross-referenced with publicly available resources (**Supplementary Table 5**). More specifically, for endotype A, prominent features consisted of: N-acetyl-3-methylhistidine and N-acetyl-1-methylhistidine – increased after muscle myofibrillar proteolysis and in renal failure(34); XIRP1 – encoding Xin, a muscle-specific actin binding protein upregulated within 12 hours of injury(35);

and MAP3K6 – a mitogen-activated protein kinase kinase involved in apoptosis signalling(36). Endotype B defining features were: complement factor H-related protein; HOXD3 – upregulation of which increases immune cell adherence by upregulating glycoprotein IIb/IIIa(37); TRIM48 – integral to interferon-γ signaling and oxidative stress-responsive cell death via apoptosis signal-regulating kinase 1(38); PPP1R3A(39); and REG3A – which encodes a bactericidal C-type lectin known commonly as pancreatitis-associated protein that, among multiple actions, alters the gut microbiome and regulates gastrointestinal inflammation(40) (**Fig 2a**). Discernable features for endotype C include GGT2 (γ-glutamyl transferase 2) – glutathione homeostasis; dopamine sulphate – a marker of increased gastrointestinal metabolism of endogenous dopamine(42); citrulline – integral to the tricarboxylic acid cycle; and SPTSSB – the rate-limiting enzyme for sphingolipid biosynthesis(43). For endotype D, the thematic features were CELA2A – pancreatic elastase 2; UDP-glucuronosyltransferase – which is associated with Gilbert-type hyperbilirubinemia(44, 45); and SLCO1B7 – a liver-specific organic anion transporter involved in bile secretion.

The link of biological function to endotype was achieved by using baseline data and computing the variable importance in projection (VIP) scores using a PLS-DA applied to the AUC+PCA 4-cluster grouping, as shown in **Figure 2a, Supplementary Figures 3 and 4.** Critically, although the PLS-DA model was built using data solely from baseline, the clusterings obtained using time=0 data produced inferior results, with Jaccard indexes never exceeding 0.75 (**Supplementary Figure 5)**. This shows that a dynamic dimension is beneficial for training an illness trajectory model.

To add further to the biological relevance of the endotypes, we also performed a compound set enrichment analysis for each endotype which highlighted processes of potential interest for each one of the endotypes (**Supplementary Figure 7**).

All participants with severe AP clustered in endotype A, despite data on clinical outcome being withheld from the model (**Figure 2b),** and this finding was statistically significant (severe vs. non-severe, Fisher's exact test, *P*=.038). Furthermore, in-hospital death was reported for two patients, and both were allocated to endotype A, confirming endotype A to be associated to individuals at higher-risk of severe illness and/or death, although this association is not proven statistically, and will require further external validation. Modified MODS score distributions for cardiovascular, renal and respiratory systems were represented and are available as **Supplementary Figure 8**. However, the etiology of AP was distributed evenly across endotypes (Fisher's exact test, *P*=1) (**Figure 2c).** Gender (Fisher's exact test, *P*= .67) or time of onset of symptoms (ANOVA, *P*=.97) were not statistically significantly associated with endotype (**Supplementary Results**). Independence between endotypes and systemic inflammatory response syndrome (SIRS) was tested for and was not rejected (SIRS vs no SIRS, Fisher's exact test, *P*=.097) (**Supplementary Figure 9**).

External validity

Having identified four endotypes in the discovery dataset, we wanted to see if that was reproducible in an independent dataset. Therefore, to externally validate the generalisability of our findings and confirm that the identified endotypes could also be detected in a distinct non-overlapping cohort, we applied our analytical methods and our results in an independent dataset of AP patients, the KAPVAL cohort (**Figure 3a and Supplementary Figures 10, Supplementary Tables 6 and 7**). Independence between reported deaths and group allocation was not rejected (Fisher's exact test, *P*=.39, **Figure 3b**). However, admission to critical care (ICU/HDU vs. ward stay only) was dependent of group allocation (Fisher's exact test, *P*<.001, **Figure 3c**).

We computed Spearman's correlation (comparison of ranks, validated using a t-distribution) for each pairwise comparison of groups from both cohorts, using metabolites that were not used to predict the group allocations (**Supplementary Figure 11, Figure 3d**). We obtained significant results when comparing groups from IMOFAP cohort to their corresponding groups in KAPVAL cohort (correlation coefficients ranged from 0.38 to 0.65, corresponding adjusted p-values <.001 for each endotype). This demonstrated underlying biological similarity between corresponding endotypes and was unlikely to be observed by chance.

We computed IGP for all endotypes, using the recruitment time point, and obtained values of 0.73, 0.51, 0.64 and 0.63 respectively for endotypes A, B, C and D. Associated FDR-corrected p-values (using permuted centroids and 10 000 iterations) were <.001 for endotypes A and D and 0.01 for endotypes B and C. This confirmed that endotypes identified in our IMOFAP cohort were also present in KAPVAL and could be identified using an early time point.

Generalisability in ARDS

We noticed an unexpected similarity between our endotypes and disease endotypes reported in a related but distinct condition, ARDS(11). Severe AP can cause ARDS, but there are multiple other causes of ARDS including sepsis, trauma, and major surgery(46). Importantly, only 6 out 56 AP-confirmed patients (no measured value available for n=1) in our discovery cohort met the Berlin definition of ARDS(47) at recruitment. Specifically, 5 of them were included in our clustering analysis, 4 (P:F ratios 130.3, 223, 273 and 286 mmHg) of which were allocated to endotype A and 1 to endotype B (P:F ration 290 mmHg). We hypothesised that the AP endotype C reflects the disease phenotype 1 ARDS endotype reported by Calfee *et al(11)*. We found 19 variables that matched the 31 variables in the ARDS study (8 physiological, 9 clinical biochemical, and 2 cytokine variables that were not used to produce the clusterings and are described, at

recruitment, in **Supplementary Table 8**). There was a significant negative correlation between endotype A with phenotype 1 in the two ARDS cohorts reported elsewhere (FDR-corrected p-values, *P*=.046 for ALVEOLI and ARMA), when considering the variables in common. Interestingly, endotype C correlated positively with the ALVEOLI (*P*=.046) and ARMA cohorts (*P*=.046). (**Figures 4a and 4b**) using the same variables. This suggests that the endotypes of AP that we report here are generalisable to another type of critical illness.

MOFA results

Finally, we highlighted a substantial overlap (Jaccard index 0.88, **Supplementary Figure 12**) between our proposed endotypes and groups highlighted using a multi-omics data integration framework, MOFA*(29),* when using AUC values.

Discussion

Disease endotyping

Disease endotypes have been commonly described using one data type, in sepsis for example(12, 13, 15). However, there are advantages to using multi-omics datasets for endotyping(48, 49). Indeed, multi-omics-based analyses have the potential to uncover biomarkers and/or processes of interest that could not be highlighted using solely one omic layer(30). It is also possible to identify commonalities and discrepancies between different data types and explore new avenues for disease study. One limitation lies in the fact that adding more layers will increase the data dimensionality. Thus, dimensionality reduction and/or variable selection are worth considering when doing endotype discovery(30). It should also be noted that some data types might not be directly comparable and should be either homogenised or analysed separately using specialised frameworks taking into account these differences.

Limitations

In this study, we have clustered AP individuals using multi-omics profiles. Four clusters were identified using AUC values projected onto a 2D-PCA space. The clusters consisted of 13, 10, 5 and 6 individuals, respectively for endotypes A, B, C and D. It should be noted that the presented solution might beneficiate from refinement using a larger cohort. Indeed, it is possible that some of the clusters could be further split into new clusters and/or new clusters identified, but not enough samples were available to do so. The identification of discriminating variables, which could ultimately be used to classify new AP patients into one of the four identified endotypes,

would also benefit from a larger cohort size. This would help in reducing the minimum number of variables required to confidently assign AP-affected individuals to one of the four clusters.

When validating our endotypes, we used the KAPVAL cohort. One time point and metabolomics measurement were available. However, we identified similar signal within the KAPVAL cohort data, when compared to our four endotypes, highlighted using time-series multi-omics data. This demonstrated that this common signal was not only driven by transcriptomics data, as it could have been expected due to the high number of gene variables.

Conclusions

Our data confirm the existence of pathobiological molecular-based endotypes in AP that do not clearly align with current clinical measures of severity or etiology. Using a novel serial evaluation approach with high-resolution multi-omics data, we systematically identify four endotypes that passed stability and biological relevance validation. These AP endotypes were validated in a second independent cohort. While it is premature to expect that identification of these specific endotypes will directly influence treatment decisions in AP today, in general, the discovery and identification of endotypes is likely to become important for future treatment choice and response to therapy in AP. Defining AP endotypes has permitted the identification of mechanisms and biomarkers of interest that could be targeted for the development of novel therapeutic strategies, and increase our understanding of AP. Allocating AP individuals to one of the endotypes could move the current paradigm towards a more personalised approach. Moreover, patients at higher risk, notably those allocated to endotype A, could be identified early in their disease trajectory, maximising their chance of a better outcome.

Moreover, we observed a statistically significant similarity with endotypes previously discovered using only clinical and cytokine measurements, at a single timepoint in two large

(n=549 and n=473) cohorts of patients with a distinct clinical syndrome, ARDS. Despite the wide differences in methodology, our approach found the same signals (**Figure 5)**.

Our unsupervised, systematic analysis re-discovered these two ARDS endotypes in an almost completely non-overlapping clinical syndrome. We conclude that these patterns may reflect generalisable components of the host response to systemic injury. Future attempts to classify subtypes within critical illness syndromes, including sepsis, trauma, AP and ARDS should recognise the generalisability of subgrouping signals.

Study approval

The IMOFAP study was approved by the Scotland A Research Ethics Committee (ref: 13/SS/0136 and amendment AM01), NHS Lothian Research & Development Project Number: 2013/0098 and amendment SA1. The University of Edinburgh and NHS Lothian ACCORD were sponsors. NHS Lothian Caldicott Guardian permission was obtained for access to identifiable patient data where needed prior to linked anonymisation. Adults without the capacity to give informed consent were recruited in accordance with the Adults with Incapacity (Scotland) Act 2000, Part 5, through their legal representative, and informed consent was sought when capacity was regained. All participants with capacity provided written informed consent. The study was registered on the UK Clinical Trials Gateway (ref: 16116). The KAPVAL cohort was approved by the East of Scotland Research Ethics Committee (ref: 15/ES/0094) and NHS Lothian Research & Development Project Number: 2015/0447/SR594, under the aegis of the Lothian NRS Human Annotated Bioresource. The University of Edinburgh and NHS Lothian ACCORD are sponsor. NHS Lothian Caldicott Guardian permission was obtained for access to identifiable patient data where needed prior to linked anonymisation.

Author contributions

Lucile P. A. Neyton, Andrea Doeschl-Wilson, Michael U. Gutmann, Iain Uings, J. Kenneth Baillie and Damian J. Mole contributed to the study conception and design and to the analysis and interpretation of the data.

Xiaozhong Zheng, Christos Skouras, Francesco V. Rao, Armel Nicolas, Craig Marshall and Lisa-Marie Wilson were involved in data acquisition and pre-processing.

The first draft of the manuscript was written by Lucile P. A. Neyton, J. Kenneth Baillie and Damian J. Mole and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

L.N. holds a Doctoral Training Program PhD studentship from the Medical Research Council. J. K.B is funded by a Wellcome Trust Intermediate Clinical Fellowship (103258/Z/13/Z) and a Wellcome-Beit Prize (103258/Z/13/A), BBSRC Institute Strategic Programme Grant to the Roslin Institute, and the UK Intensive Care Society. D.J.M. is a Senior Clinical Fellow of the Medical Research Council (MR/P008887/1). The IMOFAP study and metabolomics are part of, and funded by, the University of Edinburgh/GSK Discovery Partnership with Academia collaboration. The study was also co-funded by the Wellcome Trust through the Institutional Strategic Support Fund 3 (ISSF3).

Conflict of interest

No author has any actual or potential conflict of interest with the publication of this paper.

References

1. Steinberg W, Tenner S. Acute pancreatitis. N Engl J Med 1994;330(17):1198-1210.

2. Xiao AY, Tan ML, Wu LM, et al. Global incidence and mortality of pancreatic diseases: a systematic review, meta-analysis, and meta-regression of population-based cohort studies. Lancet Gastroenterol Hepatol 2016;1(1):45-55.

3. Peery AF, Dellon ES, Lund J, et al. Burden of gastrointestinal disease in the United States: 2012 update. Gastroenterology 2012;143(5):1179-1187.e1171-1173.

4. Kang R, Lotze MT, Zeh HJ, et al. Cell death and DAMPs in acute pancreatitis. Mol Med 2014;20:466-477.

5. Mole DJ, Webster SP, Uings I, et al. Kynurenine-3-monooxygenase inhibition prevents multiple organ failure in rodent models of acute pancreatitis. Nat Med 2016;22(2):202-209.

6. Forsmark CE, Baillie J. AGA Institute technical review on acute pancreatitis. Gastroenterology 2007;132(5):2022-2044.

7. Mole DJ, Gungabissoon U, Johnston P, et al. Identifying risk factors for progression to critical care admission and death among individuals with acute pancreatitis: a record linkage analysis of Scottish healthcare databases. BMJ Open 2016;6(6):e011474.

8. Mole DJ, McClymont KL, Lau S, et al. Discrepancy between the extent of pancreatic necrosis and multiple organ failure score in severe acute pancreatitis. World J Surg 2009;33(11):2427-2432.

9. Tenner S, Sica G, Hughes M, et al. Relationship of necrosis to organ failure in severe acute pancreatitis. Gastroenterology 1997;113(3):899-903.

10. Burnham KL, Davenport EE, Radhakrishnan J, et al. Shared and Distinct Aspects of the Sepsis Transcriptomic Response to Fecal Peritonitis and Pneumonia. Am J Respir Crit Care Med 2017;196(3):328-339.

11. Calfee CS, Delucchi K, Parsons PE, et al. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. Lancet Respir Med 2014;2(8):611-620.

12. Davenport EE, Burnham KL, Radhakrishnan J, et al. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. Lancet Respir Med 2016;4(4):259-271.

13. Scicluna BP, van Vught LA, Zwinderman AH, et al. Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study. Lancet Respir Med 2017;5(10):816-826.

14. Sweeney TE, Shidham A, Wong HR, et al. A comprehensive time-course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. Sci Transl Med 2015;7(287):287ra271.

15. Sweeney TE, Azad TD, Donato M, et al. Unsupervised Analysis of Transcriptomics in Bacterial Sepsis Across Multiple Datasets Reveals Three Robust Clusters. Crit Care Med 2018;46(6):915-925.

16. Wong HR, Cvijanovich N, Lin R, et al. Identification of pediatric septic shock subclasses based on genome-wide expression profiling. BMC Med 2009;7:34.

17. Crosby D, Bossuyt P, Brocklehurst P, et al. The MRC Framework for the Development, Design and Analysis of Stratifed Medicine Research. Draft 2017.

18. Russell CD, Baillie JK. Treatable traits and therapeutic targets: goals for systems biology in infectious disease. Current Opinion in Systems Biology 2017.

19. Skouras C, Zheng X, Binnie M, et al. Increased levels of 3-hydroxykynurenine parallel disease severity in human acute pancreatitis. Sci Rep 2016;6:33951.

20. Banks PA, Bollen TL, Dervenis C, et al. Classification of acute pancreatitis--2012: revision of the Atlanta classification and definitions by international consensus. Gut 2013;62(1):102-111.

21. Lough G, Kyriazakis I, Bergmann S, et al. Health trajectories reveal the dynamic contributions of host genetic resistance and tolerance to infection outcome. Proc Biol Sci 2015;282(1819).

22. Giorgino T. Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package. 2009 2009;31(7):24.

23. Hall DP, MacCormick IJ, Phythian-Adams AT, et al. Network analysis reveals distinct clinical syndromes underlying acute mountain sickness. PLoS One 2014;9(1):e81229.

24. Hennig C. Dissolution point and isolation robustness: Robustness criteria for general cluster analysis methods. Journal of Multivariate Analysis 2008;99(6):1154-1176.

25. Kanehisa M, Furumichi M, Tanabe M, et al. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 2017;45(D1):D353-D361.

26. Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. Nature 2014;507(7493):455-461.

27. Lizio M, Harshbarger J, Shimoji H, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol 2015;16:22.

28. Kapp AV, Tibshirani R. Are clusters found in one dataset present in another dataset? Biostatistics 2007;8(1):9-31.

29. Argelaguet R, Velten B, Arnol D, et al. Multi-Omics Factor Analysis-a framework for unsupervised integration of multi-omics data sets. Mol Syst Biol 2018;14(6):e8124.

30. Pinu FR, Beale DJ, Paten AM, et al. Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community. Metabolites 2019;9(4).

31. Strauss T, von Maltitz MJ. Generalising Ward's Method for Use with Manhattan Distances. PLoS One 2017;12(1):e0168288.

32. Miyamoto S, Abe R, Endo Y, et al. Ward method of hierarchical clustering for non-Euclidean similarity measures. In: 7th International Conference of Soft Computing and Pattern Recognition (SoCPaR); 2015; 2015.

33. Szekely GJ, Rizzo ML. Hierarchical Clustering via Joint Between-Within Distances: Extending Ward's Minimum Variance Method. Journal of Classification 2005;22(2):151-183.

34. Mondry A. Antiglucocorticoid RU38486 reduces net protein catabolism in experimental acute renal failure. BMC Nephrol 2005;6:2.

35. Hawke TJ, Atkinson DJ, Kanatous SB, et al. Xin, an actin binding protein, is expressed within muscle satellite cells and newly regenerated skeletal muscle fibers. Am J Physiol Cell Physiol 2007;293(5):C1636-1644.

36. Okazaki T, Higuchi M, Takeda K, et al. The ASK family kinases differentially mediate induction of type I interferon and apoptosis during the antiviral response. Sci Signal 2015;8(388):ra78.

37. Taniguchi Y, Komatsu N, Moriuchi T. Overexpression of the HOX4A (HOXD3) homeobox gene in human erythroleukemia HEL cells results in altered adhesive properties. Blood 1995;85(10):2786-2794.

38. Hirata Y, Katagiri K, Nagaoka K, et al. TRIM48 Promotes ASK1 Activation and Cell Death through Ubiquitination-Dependent Degradation of the ASK1-Negative Regulator PRMT1. Cell Rep 2017;21(9):2447-2457.

39. Delibegovic M, Armstrong CG, Dobbie L, et al. Disruption of the striated muscle glycogen targeting subunit PPP1R3A of protein phosphatase 1 leads to increased weight gain, fat deposition, and development of insulin resistance. Diabetes 2003;52(3):596-604.

40. Darnaud M, Dos Santos A, Gonzalez P, et al. Enteric Delivery of Regenerating Family Member 3 alpha Alters the Intestinal Microbiota and Controls Inflammation in Mice With Colitis. Gastroenterology 2018;154(4):1009-1023.e1014.

41. Jetter A, Kinzig-Schippers M, Illauer M, et al. Phenotyping of N-acetyltransferase type 2 by caffeine from uncontrolled dietary exposure. European Journal of Clinical Pharmacology 2004;60(1):17-21.

42. Goldstein DS, Swoboda KJ, Miles JM, et al. Sources and physiological significance of plasma dopamine sulfate. J Clin Endocrinol Metab 1999;84(7):2523-2531.

43. Hanada K. Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. Biochim Biophys Acta 2003;1632(1-3):16-30.

44. Lankisch TO, Behrens G, Ehmer U, et al. Gilbert's syndrome and hyperbilirubinemia in protease inhibitor therapy--an extended haplotype of genetic variants increases risk in indinavir treatment. J Hepatol 2009;50(5):1010-1018.

45. Lankisch TO, Gillman TC, Erichsen TJ, et al. Aryl hydrocarbon receptor-mediated regulation of the human estrogen and bile acid UDP-glucuronosyltransferase 1A3 gene. Arch Toxicol 2008;82(9):573-582.

46. Thompson BT, Chambers RC, Liu KD. Acute Respiratory Distress Syndrome. N Engl J Med 2017;377(19):1904-1905.

47. Ranieri VM, Rubenfeld GD, Thompson BT, et al. Acute respiratory distress syndrome: the Berlin Definition. JAMA 2012;307(23):2526-2533.

48. Lefaudeux D, De Meulder B, Loza MJ, et al. U-BIOPRED clinical adult asthma clusters linked to a subset of sputum omics. J Allergy Clin Immunol 2017;139(6):1797-1807.

49. Obeso D, Mera-Berriatua L, Rodríguez-Coira J, et al. Multi-omics analysis points to altered platelet functions in severe food-associated respiratory allergy. Allergy 2018;73(11):2137-2149.

Figure legends

Figure 1.

a. Flow diagram for the IMOFAP cohort. Study flow chart for patient samples and data from the IMOFAP study included in the analysis showing filtering process and reasons for exclusion. After filtering, remaining samples (n=54) could be used for further analyses, however, some presented a single time point (n=20) and thus were not included in analyses requiring several samples over a time course.

b. Sample and data time points. Samples used to generate data types and corresponding time points are represented. Time point 0 corresponds to recruitment time point (some clinical data was collected when a patient was admitted to the hospital and is represented at t=-5 hours on the figure). Dashed lines show the median time from admission to hospital to intensive care admission in those who required it (12 hours) and median time from admission to hospital to death for AP fatalities (82 hours)(7).

c. AP endotypes discovery process. Data analysis overview using the 34 pre-selected IMOFAP individuals (individuals with less than 2 time points for one type of omics were not included in the analysis, n=20). Each row represents results for one of the three time-series-based method. Within the clustering section, hierarchical trees for each time series-based clustering method are presented along with the optimal solution (the optimal number of clusters and respective sample allocations being represented by branches colour), patient identifiers are reported on the x axis and distances are presented on the y axis. Stability measures are reported for all displayed solution (average Jaccard index) and a summary of the number of compound sets significantly enriched is shown for each category (respectively FANTOM5 results, gene-based results and metabolic compound

results) in the evaluation section. Results based on a single time point (using Euclidean distances) were not presented here as the clusters obtained using these presented poor structure (generally a main group in which the majority of individuals clustered and a number of single-patient groups). For the selected solution (Area-under-the-curve combined with PCA), findings were reproduced using one AP cohort and compared to previously defined ARDS endotypes, as illustrated in the reproducibility section.

Figure 2.

a. AP endotypes. Heatmap of normalised and scaled variable values of the top 10 variables, averaged for all patients belonging to each endotype, for all four subgroups. Compound names, on the y axis, are reported in grey italic when corresponding to genes. Other compounds refer to metabolites. Group labels are reported on the bottom x and on the left y axis to highlight which variables correspond to which group. For visualisation purposes row values were scaled between 0 and 1. Colours are representative of the range of values observed.

b. Clinical severity distribution. For comparison purposes, distribution of clinical severity categorised by mMODS score within each endotype. For each one of the four endotypes, the number of patients in each severity category (mild, moderate or severe) is shown using a colour code. The x axis was used to represent the identified endotypes and the y axis to show the number of patients, per endotype, falling in each one of the severity categories.

c. Etiology distribution. Distribution of etiology within each endotype. For each endotype, the number of patients per unique etiology is shown. The x axis was used to represent the identified endotypes and the y axis to show the number of patients, per endotype, falling in each one of the etiology categories.

Figure 3.

a. Allocation of KAPVAL samples to IMOFAP-based endotypes. Schematics representing the process to assign KAPVAL individuals to endotypes identified in IMOFAP cohort. This is for illustrative purposes only, variables and values represented do not correspond to the models created or the data used to build them. After identifying metabolites in common between the IMOFAP and KAPVAL cohorts, one PLS-DA model per endotype is built, choosing a number of metabolites to include to maximise the prediction accuracy. The four obtained models are then used to predict allocations for all KAPVAL samples, returning a value for each. A sample will be allocated to a group given the model for which the largest PLS-DA predicted value was obtained (corresponding to the closest match for a given sample).

b. In-hospital mortality for the KAPVAL cohort. Using predicted allocations for KAPVAL samples, distribution of in-hospital mortality within each endotype. 1 refers to death and 0 to no death. The x axis represents the allocation endotypes for KAPVAL samples and the y axis the number of patients, per allocation endotype, falling in each one of the mortality categories.

c. Care level for the KAPVAL cohort. Distribution of care level for KAPVAL individuals given endotype allocation. Three categories of care level were used (ward, HDU for high-dependency unit and ICU for intensive care unit) and correspond to the three represented colours. The x axis shows the allocation endotypes for KAPVAL samples and the y axis the number of patients, per allocation endotype, falling in each one of the care level categories.

d. Internal validation of AP endotypes. Correlation matrix representing Spearman's correlation results for pairwise comparisons between training (IMOFAP) and testing (KAPVAL) data. Spearman's correlation coefficients were computed using average values per variable per endotype

for AUC-PCA-based four clusters in IMOFAP (x axis) and corresponding predicted KAPVAL endotypes (with allocated samples, on the y axis). Correlation coefficients are represented using a colour code, as described by the colour bar on the figure's right-hand side. FDR-adjusted p-values are reported for each cell.

Figure 4.

a. ALVEOLI/ARMA and IMOFAP cohorts comparison for endotypes A and C. Variable values available as part of the IMOFAP study that occur in common with those reported in the ARDS study of Calfee *et al* are compared. Ranks of normalised values (Z-scores) are represented for endotypes A and C against ranks from Calfee *et al* results for the ALVEOLI cohort (upper panel) and the ARMA cohort (lower panel). Linear trends are plotted along Spearman's correlation coefficients (using reported variable ranks) and corresponding FDR-corrected p-values.

b. Reproducibility of AP endotypes in two ARDS cohorts. Comparisons between ARDS cohorts (ALVEOLI and ARMA) and IMOFAP endotypes (A, B, C and D). Using matched variables, Spearman's correlation coefficients between ranked variables from Calfee *et al* and average Z-scores for the four AP endotypes were computed for all pairwise comparisons. Colours represent correlation coefficients values and FDR-corrected p-values are reported within each cell for all pairwise comparisons.

Figure 5.

Endotype model. Systematic inflammatory endotypes model. Highlighted endotypes in AP patients are represented alongside their overlap with ARDS endotypes described in Calfee *et al*. The top part of the figure shows pancreas injury along with their potential initiators and

corresponding endotypes. This is reproduced in the figure bottom part, for ARDS, representing lung injury and some of their initiators and inflammatory endotypes.

Figure 1

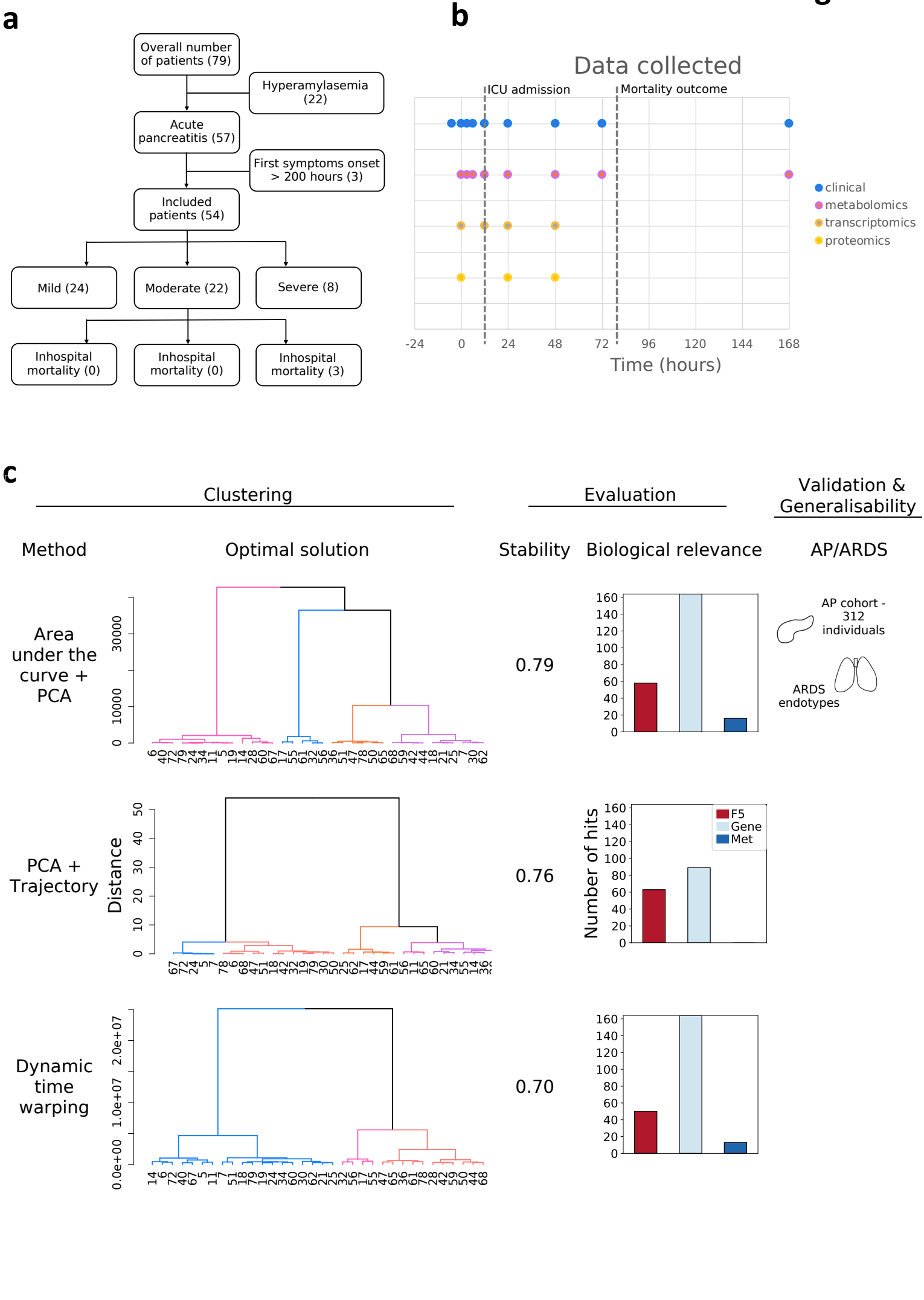
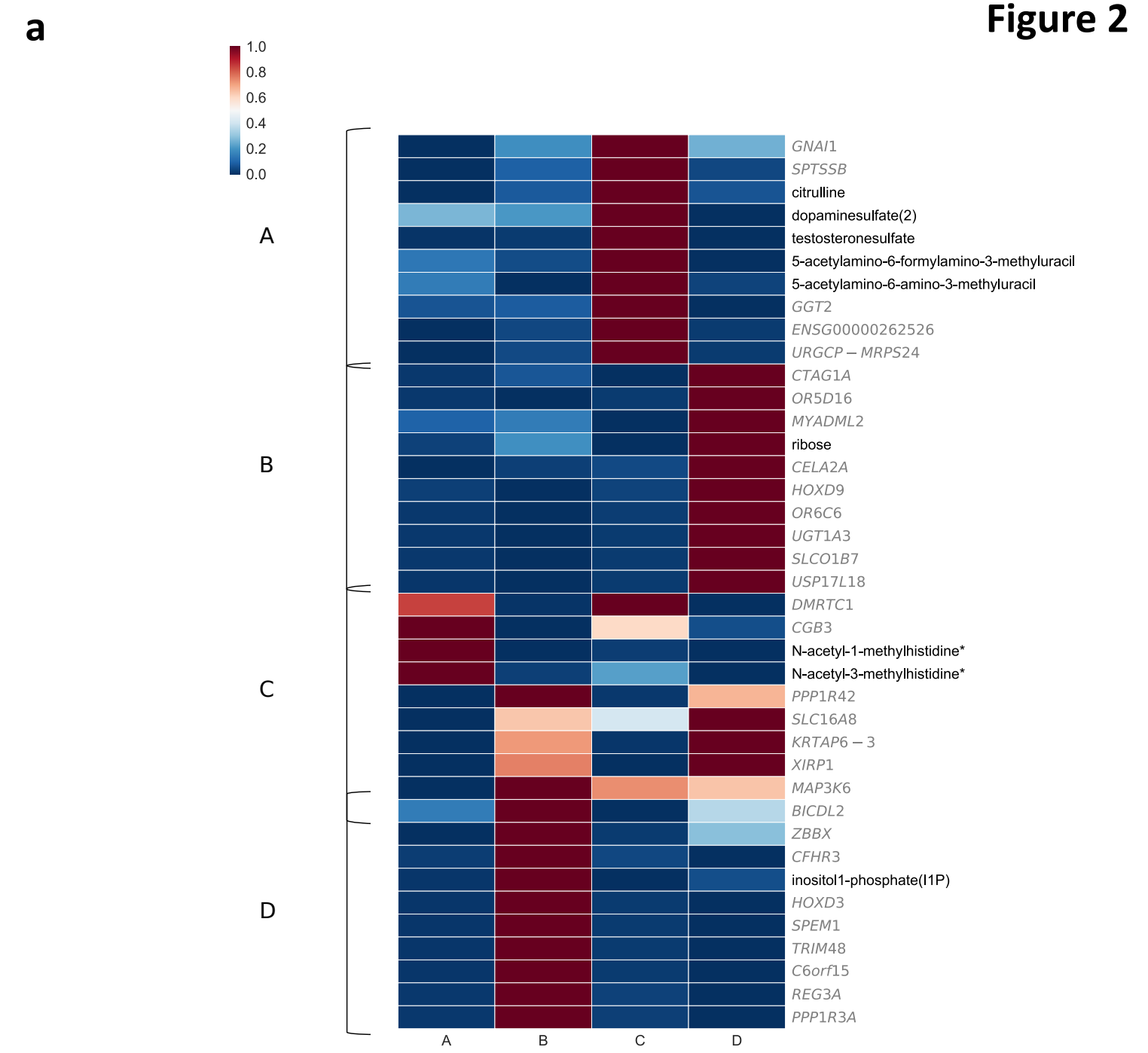
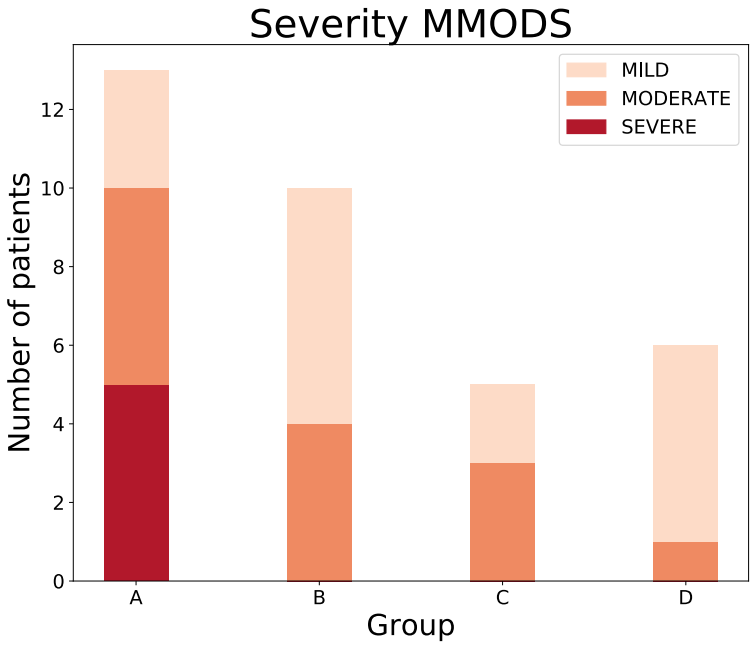

a
Overall number of patients (79)
Hyperamylasemia (22)
Acute pancreatitis (57)
First symptoms onset > 200 hours (3)
Included patients (54)
Mild (24)
Moderate (22)
Severe (8)
Inhospital mortality (0)
Inhospital mortality (0)
Inhospital mortality (3)
b
Data collected
ICU admission
Mortality outcome
clinical
metabolomics
transcriptomics
proteomics
-24
0
24
48
72
96
120
144
168
Time (hours)
c
Clustering
Evaluation
Validation & Generalisability
Method
Optimal solution
Stability
Biological relevance
AP/ARDS
Area under the curve + PCA
0.79
AP cohort - 312 individuals
ARDS endotypes
PCA + Trajectory
Distance
0.76
Number of hits
F5
Gene
Met
Dynamic time warping
0.70

Figure 2


a
1.0
0.8
0.6
0.4
0.2
0.0
A
B
C
D
GNAI1
SPTSSB
citrulline
dopaminesulfate(2)
testosteronesulfate
5-acetylamino-6-formylamino-3-methyluracil
5-acetylamino-6-amino-3-methyluracil
GGT2
ENSG00000262526
URGCP – MRPS24
CTAG1A
OR5D16
MYADML2
ribose
CELA2A
HOXD9
OR6C6
UGT1A3
SLCO1B7
USP17L18
DMRTC1
CGB3
N-acetyl-1-methylhistidine*
N-acetyl-3-methylhistidine*
PPP1R42
SLC16A8
KRTAP6 – 3
XIRP1
MAP3K6
BICDL2
ZBBX
CFHR3
inositol1-phosphate(I1P)
HOXD3
SPEM1
TRIM48
C6orf15
REG3A
PPP1R3A
A
B
C
D


b
Severity MMODS
MILD
MODERATE
SEVERE
Number of patients
0
2
4
6
8
10
12
A
B
C
D
Group

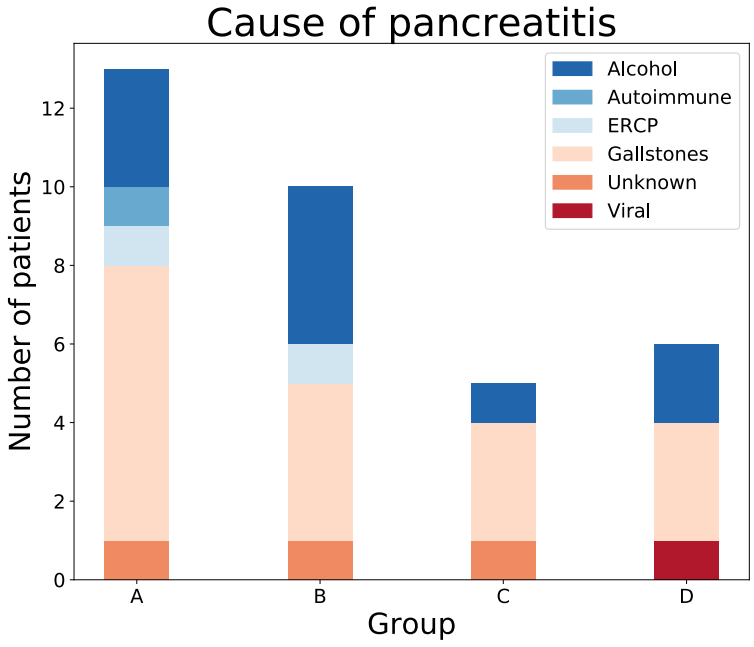

c
Cause of pancreatitis
Alcohol
Autoimmune
ERCP
Gallstones
Unknown
Viral
Number of patients
0
2
4
6
8
10
12
A
B
C
D
Group

Figure 3

a

ID	group	allantoin	citrulline	codeine	orotate	ribose	serine	xylose
5	3	-0.3062	-0.2266	-0.1218	-0.1779	-0.4771	-0.2121	-0.1371
14	3	-0.5530	-0.0453	0.9011	-0.1534	0.5191	-1.7056	0.1578
2	1	-0.3092	-0.1001	-0.1798	-0.3121	0.9146	0.2981	-0.2759
60	2	0.5028	-0.0453	-0.1798	-0.1056	-0.4270	-0.7262	0.1052
76	4	-0.2109	-0.2595	-0.1799	-0.0569	-0.3210	0.4076	-0.2451
11	1	-0.2870	-1.5231	-0.1537	-0.0373	-0.1629	-1.1724	0.0257

IMOFAP

Baseline metabolomics

Endotype models

y: *1 for current group*
-1 for any other group

Build models

A: Selected predictors: orotate and xylose

B: Selected predictors: codeine

C: Selected predictors: allantoin and citrulline

D: Selected predictors: xylose

Apply models

Patient 005

$predicted\ value\ vg_1$ $predicted\ value\ vg_2$ $predicted\ value\ vg_3$ $predicted\ value\ vg_4$

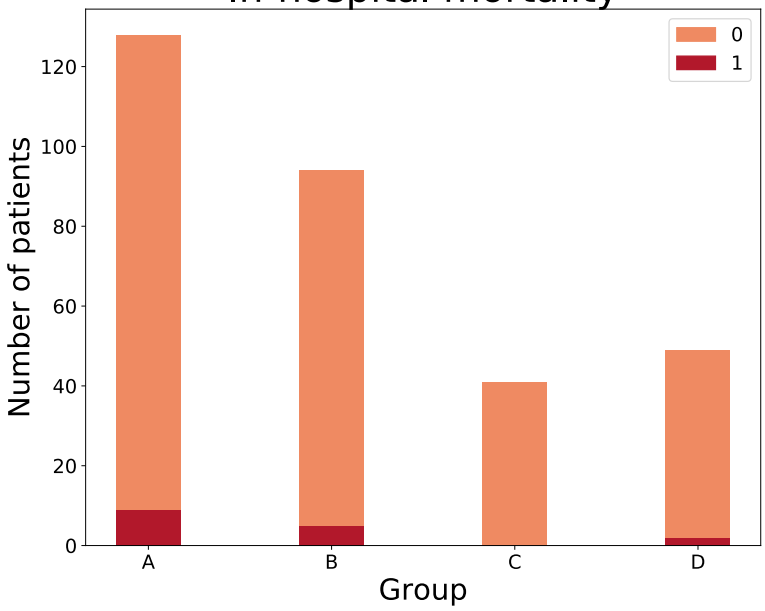
Patient 005 allocated based on largest predicted value

KAPVAL

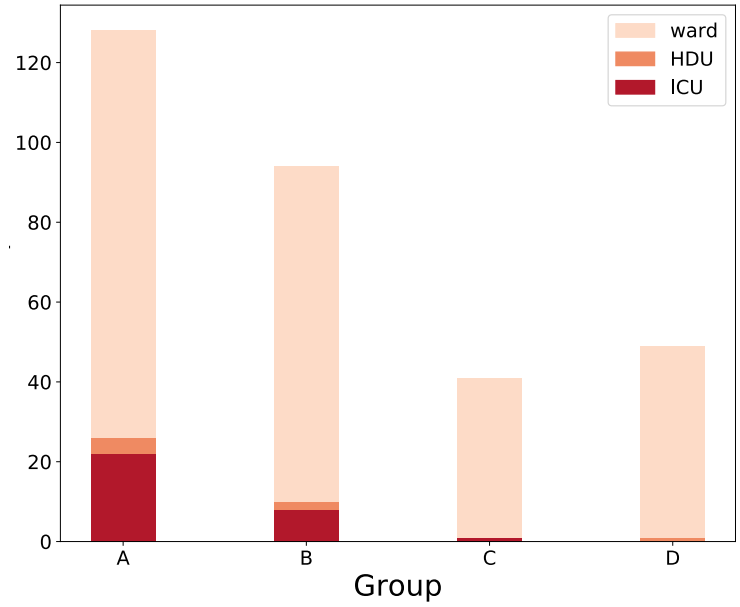
Patient 005

ID	allantoin	citrulline	codeine	orotate	xylose
005	-0.306	-0.227	-0.122	-0.178	-0.137

b



c



d

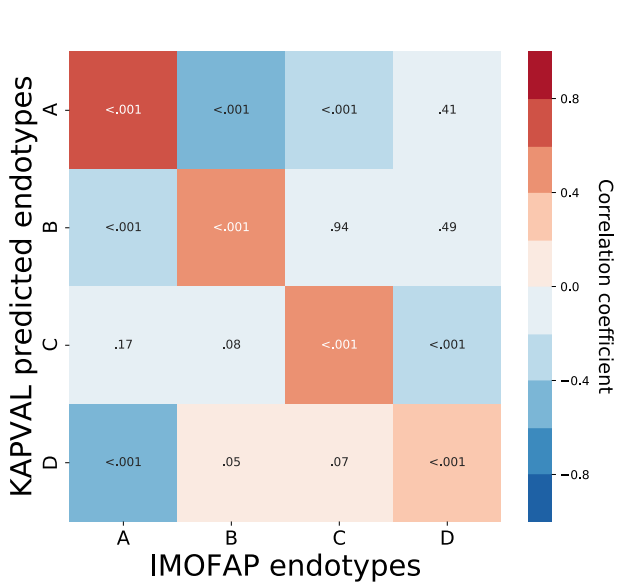
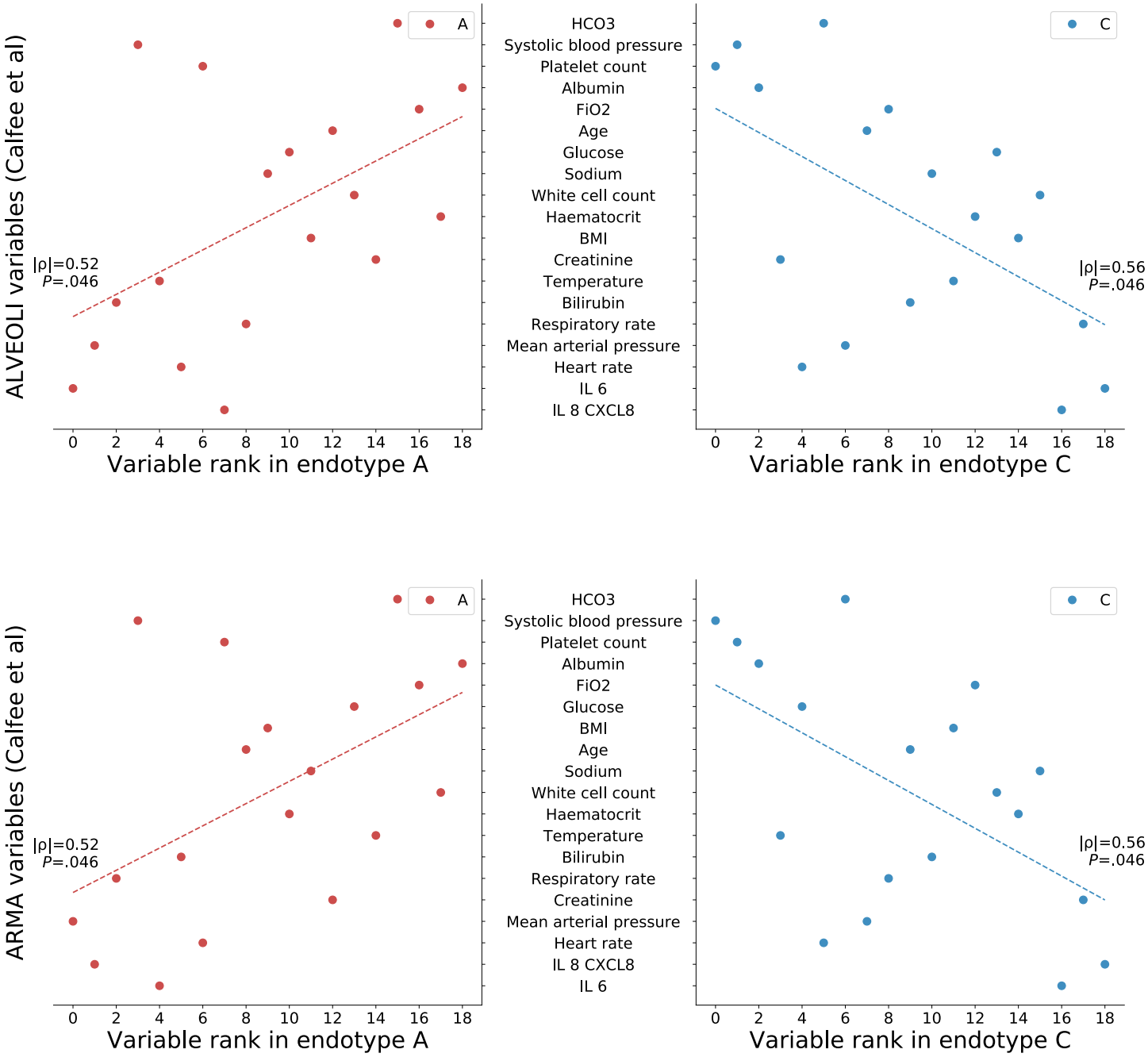


Figure 4

a


ALVEOLI variables (Calfee et al)
A
|ρ|=0.52
P=.046
Variable rank in endotype A
HCO3
Systolic blood pressure
Platelet count
Albumin
FiO2
Age
Glucose
Sodium
White cell count
Haematocrit
BMI
Creatinine
Temperature
Bilirubin
Respiratory rate
Mean arterial pressure
Heart rate
IL 6
IL 8 CXCL8
C
|ρ|=0.56
P=.046
Variable rank in endotype C
ARMA variables (Calfee et al)
A
|ρ|=0.52
P=.046
Variable rank in endotype A
HCO3
Systolic blood pressure
Platelet count
Albumin
FiO2
Glucose
BMI
Age
Sodium
White cell count
Haematocrit
Temperature
Bilirubin
Respiratory rate
Creatinine
Mean arterial pressure
Heart rate
IL 8 CXCL8
IL 6
C
|ρ|=0.56
P=.046
Variable rank in endotype C

b

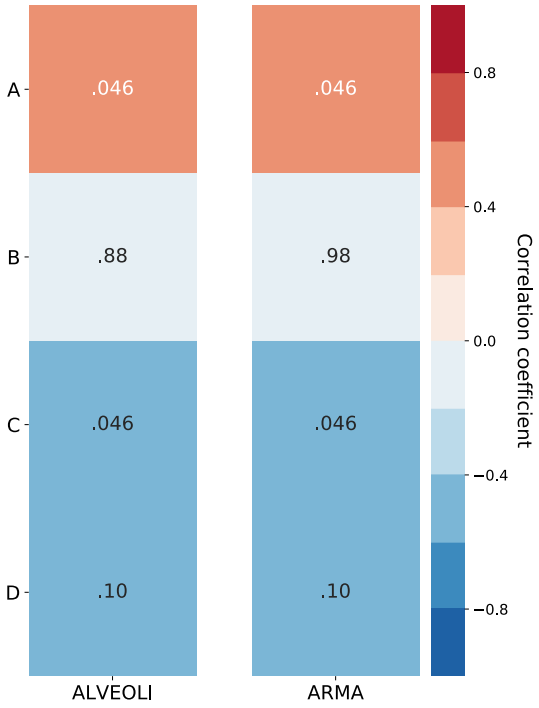
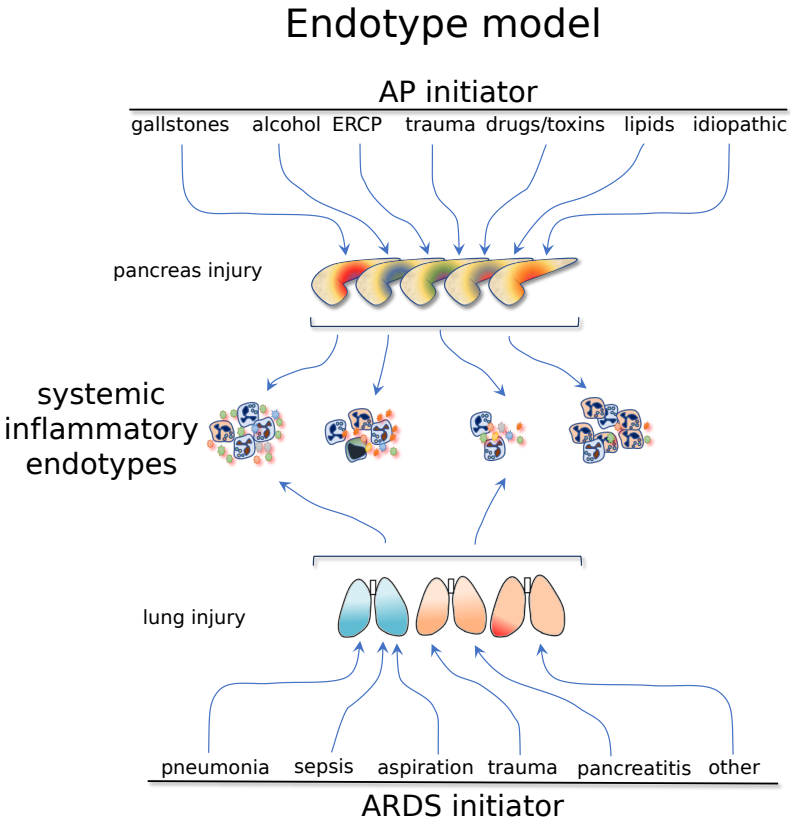

A
B
C
D
.046
.88
.046
.10
.046
.98
.046
.10
ALVEOLI
ARMA
Correlation coefficient
0.8
0.4
0.0
−0.4
−0.8

Figure 5


Endotype model
AP initiator
gallstones
alcohol
ERCP
trauma
drugs/toxins
lipids
idiopathic
pancreas injury
systemic inflammatory endotypes
lung injury
pneumonia
sepsis
aspiration
trauma
pancreatitis
other
ARDS initiator

Supplementary Material

to accompany manuscript: Molecular patterns in acute pancreatitis reflect generalisable endotypes of the host response to systemic injury in humans, by Lucile P. A. Neyton, et al.

Supplementary Materials and Methods

Cohorts

IMOFAP study details

In brief, consecutive emergency attendees to hospital with sudden onset abdominal pain with nausea and/or vomiting, and a serum amylase value greater than 100 iU/L were identified using an automated laboratory alert system with clinical verification, at all times of day or night for a period of three months. Confirmation of the diagnosis of AP according to the revised Atlanta criteria(1) was made after recruitment in 57 of 79 recruited patients, all of them having developed AP before recruitment. Samples were aliquoted into specific tubes containing appropriate preservation solution for subsequent DNA and RNA extraction, or serum and plasma extracts were prepared after centrifugation and snap frozen and the cold chain maintained until acquisition.

Although every effort was made to capture all time points, this was not always possible, and on occasion (for example, at the request of the patient to omit a short interval repeat venepuncture), time points were omitted.

KAPVAL study details

The KAPVAL cohort is a fully-linked anonymised biosample cohort which is made up of samples and data from all patients presenting to the Royal Infirmary of Edinburgh with a serum amylase level > 300 iU/L (3-fold above the upper limit of the reference range for our laboratory), clinically annotated and coded ad compliant with the revised Atlanta criteria(1) for the diagnosis of AP. An aliquot of gel-clot activator serum is retained and stored at -80 °C for all patient samples that have an elevated serum amylase. Using a linked anonymisation code, to which the investigators are blinded, the diagnosis of acute pancreatitis was confirmed by trained members of a specialist data collection team, using clinical and laboratory data obtained from the individuals electronic health record. The variables used in the analysis include age, gender, date and time of admission and discharge, diagnostic codes (3 levels), standard clinical biochemistry and clinical haematology tests, level of critical care, duration of critical care and mortality. Serum samples were used as described.

Data acquisition

Transcriptomics data acquisition

2.5 ml of peripheral venous blood was collected into the PAXgene Blood RNA Tube (BD Biosciences) following the manufacturer's instructions and stored at -80 °C until used. Total RNA was extracted and purified using the PAXgene Blood miRNA Kit (QIAGEN). The RNA integrity of total RNA samples was assessed using the Agilent 2100 Bioanalyzer. The mRNA in a total RNA sample was converted into a library of template molecules of known strand origin using the reagents provided in an Illumina® TruSeq® Stranded mRNA library prep workflow. The subsequent sequence data was obtained using Illumina HiSeq 4000 75PE system.

Proteomics data acquisition

Serum was obtained from peripheral venous blood by centrifugation and stored at -80 °C until used. Sera were subjected to depletion of abundant serum proteins using Proteome Purify 12 Human Serum Protein Immunodepletion Resin (R&D Systems). Denaturing was followed by alkylation with N-ethylmaleimide and acetone precipitation. Digestion used lysyl endopeptidase (LysC) and trypsin before labelling with 10plex TMT reagents (Thermo Fisher Scientific). TMT-labelled peptides were fractionated into 4 fractions each by High-pH Reverse Phase chromatography then each fraction analysed by RPLC-MS/MS/MS (70 min linear gradient) on a Fusion Tribrid Orbitrap operating in Data Dependent Acquisition mode (MultiNotch Simultaneous Precursor Selection method; MS1: profile mode, Orbitrap resolution 120k, 375-1550 m/z, AGC target 200,000, 50 ms max. injection time, RF lens 60%; MS2: centroid mode, IonTrap, 12 dependent scans, 1.2 Th isolation window, charge states 2-7, 60 s dynamic exclusion, CID fragmentation (35%, activation Q 0.25), AGC target 10,000, 70 ms max. injection time; MS3: profile mode, 5 precursors, 2 Th isolation window, Orbitrap resolution 50k, 100-500 m/z, AGC target 50,000, 105 ms max. injection time, HCD fragmentation (60%)). Control samples were used as internal cross-channel controls in different TMT samples and in different TMT channels to avoid any specific bias. Raw files were searched with MaxQuant (version 1.5.7.4) against a human proteome obtained from UniProt, with the match-between-runs option selected to allow for transfer of peptide identifications between files.

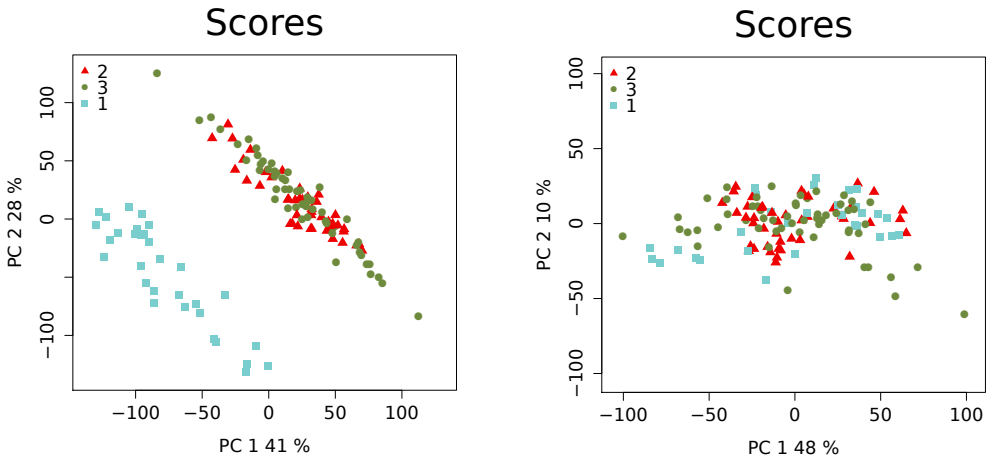
Metabolomics data acquisition

Serum was obtained from peripheral venous blood by centrifugation and stored at -80 °C until used. Aliquots of sera were shipped on dry ice to Metabolon Inc., 617 Davis Drive, Suite 400, Durham, NC 27713 USA. Serum samples underwent automated protein depletion using methanol (MicroLab STAR® system) followed by four fraction analysis by UPLC-MS/MS with positive ion mode electrospray ionization, UPLC-MS/MS with negative ion mode electrospray ionization, LC polar platform and, GC-MS. QA/QC steps included: a pooled matrix sample as a technical replicate throughout, extracted water samples as process blanks, and a bespoke cocktail of QC standards spiked into every sample for instrument performance monitoring and chromatographic alignment. Instrument variability was determined by calculating the median relative standard deviation (RSD) for the standards. Overall process variability was determined by calculating the median RSD for all endogenous metabolites (i.e. non-instrument standards) present in 100% of the pooled matrix samples. Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy (UPLC-MS/MS): The LC/MS portion of the platform used a Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyser operated at 35,000 mass resolution. Gas Chromatography-Mass Spectroscopy (GC-MS): The samples for GC-MS was derivatized under dried nitrogen using bistrimethyl-silyltrifluoroacetamide and separated on a 5% diphenyl / 95% dimethyl polysiloxane fused silica column (20 m x 0.18 mm ID; 0.18 um film thickness) with helium as carrier gas and a temperature ramp from 60° to 340°C in a 17.5 min period and analysed on a Thermo-Finnigan Trace DSQ fast-scanning single-quadrupole mass spectrometer using electron impact ionization (EI) and operated at unit mass resolving power. Data Extraction and Compound Identification: Raw data were extracted, peak-identified and QC processed using Metabolon's hardware and software and compounds were identified by comparison to library entries of purified standards or recurrent unknown entities. Metabolite Quantification and Data Normalization: Peaks were quantified using area-under-the-curve. Where runs spanned multiple days, a data normalization step was performed to correct variation resulting from instrument inter-day tuning differences.

Data pre-processing

Transcriptomics data pre-processing

RNA-Seq data consisted of 75bp paired-end Illumina reads stored as FASTQ files. One batch was carried out using polyA selection and the other using rRNA depletion**.** Reads were filtered based on QC results (FASTQC v0.11.2) to trim low-quality ends using a phred threshold of 20 and discard resulting reads shorter than 25bp. Read alignment was performed against the genome assembly hg38 using STAR(2) (v2.5.0a). Counts were generated as a proxy for gene expression by assigning previously aligned reads to exons using the tool featureCounts(3) (v1.5.2). hg38 genome was used as the reference genome. The difference in RNA sequencing (library preparation) was accounted for using a protein-coding only filter, a batch removal algorithm (using the ARSyNseq function from NOISeq R library(4)) and a normalisation step (FPKM). Finally, the normalised counts were transformed into Z-scores. This allowed a comparison across samples. PCA plots of the counts before and after batch effect removal are available in **Supplementary Figure A**.



Supplementary Figure A. Batch effect correction for RNA-Seq data. PCA plots before and after batch effect removal. RNA-Seq counts values obtained using featureCounts, for coding genes only, are represented in the left figure. The same counts, after batch effect correction and FPKM normalisation are represented in the right figure. We chose to represent the top two principal components (PC1 on the x axis and PC2 on the y axis) in terms of represented variance (as reported on the corresponding axes). Batches are represented using defined colours and shapes, as presented in the legend element.

Proteomics data pre-processing

After raw data acquisition and initial processing to generate intensity-based values, any protein species with 90% or more missing values was discarded. Scaling and linear imputation

were applied using the minimum value for each compound, as any missing value would suggest a value below the detection limit. As samples from similar batches grouped together when performing the clustering step, we corrected for it with ComBat, using a parametric empirical Bayes framework, to remove the irrelevant variation between samples due to the different runs carried out. Measurements were then transformed into Z-scores.

Metabolomics data pre-processing

Metabolomics data consisted of an abundance list (raw ion counts). Data were pre-processed using a similar pipeline as the one used for the proteomics data but did not require a batch effect removal step. Data from the KAPVAL cohort was pre-processed in the same way.

Data analysis

Analyses were carried out using Python (version 3.5) and R (version 3.3.2). Libraries used include dtwclust in R and numpy, pandas, rpy2, scipy, sklearn, statsmodels in Python.

The total starting data set consisted of the relative expression of a normalised set of 19766 genes, integrated with 1383 protein and 686 metabolite abundances. After data pre-processing, the dataset consisted of 651 metabolites, 371 proteins and 19766 genes.

Distance metrics and clustering

Area Under the Curve and PCA

Each time series was summarised by computing its AUC value, a single value representing the cumulative magnitude of response over time and computed using the trapezoidal rule (based on linear interpolation by interval). Consequently, values obtained for each variable could be treated as independent from others. We normalized the values based on the time difference between the first and last included time points. Using this new dataset we projected the values in a principal component analysis space where, selecting the first two components, we computed Euclidean distances between the different individuals. The Euclidean distances were weighted according to the proportion of variance explained by each principal component. PCA is a method of choice when encountering high dimensionality data (because of its dimensionality reduction ability), as much for data visualization as for data analysis, and is hypothesis free(5). Indeed, PCA is only sensitive to the correlation structure in the data and does not make specific assumptions related to the stratification of the input data.

Trajectory through PCA space

As described in another study(6), trajectories of patients through selected components can be helpful when clustering patients. Here we projected all time points of each individual onto a two-dimensional PCA space and looked at their evolution through this newly defined space. To characterize the trajectory of an individual through this space we considered the direction taken between each pair of time points for these particular patients. We coded this direction with a value of 1, 2, 3 or 4 depending on the direction taken when dividing the space into four quadrants. This was repeated for all patients and eventually a vector of directions was obtained for each patient.

The Hamming distance was used to compare the vectors. When taking two vectors of the same length, the Hamming distance is computed by counting, in an element-wise fashion, the number of different elements. The values were then used as dissimilarity measures between patients. This allowed to combine the advantages of PCA and trajectory analysis.

Dynamic time warping

Computation of distances between patients using dynamic time warping was carried out using the dtwclust package in R. The algorithm then considers each pair of samples. More specifically, for each variable, a matrix is generated and reports the difference in magnitude between all possible pairs of time points. One matrix of dimensions (time series 1 length*time series 2 length) is obtained for each variable. A summary matrix is generated for each patient by summing the individual matrices associated to each variable, element-by-element. The warping consists of finding a trajectory in that matrix that will minimize the distance between the two series. The process will start from the matrix element corresponding to the first points of the two series being compared (in other terms the first element of the summary matrix) and will end when reaching the last series points (corresponding to the last element of the matrix). As this is performed on a summarized matrix, the selected warping represents a consensus alignment minimizing the summed differences in magnitude between the two multivariate time-series. From this warping a final distance is calculated and can be used as a dissimilarity value. We performed linear imputation on time series when time points were not equally spaced or missing. It was preferred as it allowed a fairer comparison when dealing with time series of different lengths/sampling pattern.

Clustering

Clustering, allowing to group individuals from a cohort, was performed using hierarchical clustering and Ward's method. Ward's method is an agglomerative method and works by minimizing distances within each group. Although it is usually used for Euclidean related distances (which is not the case for all the presented methods) it has been used successfully for other types of distance(7-9) and produced the best results here when compared with others (in terms of validation results). The number of clusters was chosen according to the stability of each solution. For all number of clusters ranging from two to twenty we assessed this using bootstrapping combined with the Jaccard index. For a chosen number of clusters *k* we replaced individuals from the initial cohort to create a new input dataset. It was then used to re-perform the clustering 100 times. Each one of the new results obtained was compared to the initial solution given *k* clusters. Each cluster from the *k* generated clusters was compared to the most similar cluster of the initial solution. The Jaccard index computed the overlap between the two and this was averaged for all matched groups that were part of a solution. After 100 repetitions we obtained a value for each *k* number of clusters by taking the average of the averaged Jaccard indexes and the solution with the highest value was chosen as the best one. Twenty was chosen as the maximum number as any greater value would have resulted in many clusters composed of one or very few individuals. This was not desirable as very little information could have been drawn from it. Additionally, we chose not to select for further analyses any solution with one group or more presenting less than three individuals.

Validation

Number of clusters and stability

For all number of clusters ranging from two to twenty we assessed stability using bootstrapping combined with the Jaccard index, as defined in the previous section.

Biological plausibility

Compound identifiers were converted when necessary using R package biomaRt. Data used consisted of KEGG(10) pathways extracted from R package GAGE(11) when analysing gene and protein data and from MetaboAnalystR(12) when analysing metabolites. As four groups were identified, the aim was to assess whether or not a subset of compounds (corresponding to a pathway) was associated with the group label. For each pathway this was tested using generalised linear models. A model was fitted to the data for each compound of a corresponding pathway with the group label being the fixed effect and the response variable

being the values associated to this gene. To assess the effect of the group on the values, we performed a likelihood test to compare the newly created model to a null model, returning a single p-value. Using Stouffer's method to combine p-values, we computed a single p-value per pathway. For elements of a pathway, individual p-values were given a weight corresponding to the inverse of the total number of pathways a gene was involved in. This, this prevented overlapping pathways from biasing our results. R function anova with test="LRT" was used to do the tests. As the gene sets tested for enrichment were the same for each one of the three methods tested to cluster individuals, p-values were used to quantify the biological relevance of a clustering. Each compound set with an associated value under the threshold of .05 was counted as differentially expressed and counts obtained for each method were compared to determine the most biologically relevant result. FANTOM5(13, 14) co-expression clusters were used to compare cell type gene signatures with our groups and thus allowing the discovery of closely involved cell types. The same strategy was applied to determine if a compound set was significantly enriched or not, using the same 5% cut-off. Results were used as a way to quantify the biological pertinence, have an overall look at the results and select clustering solutions.

Enrichment analysis

Partial Least Square Discriminant Analysis (PLS-DA) is a classification algorithm. PLS-DA models were used to select variables for the enrichment analysis. Given group labels it will project the data onto a new space, given a number of components selected by the user, and then rotate the components to maximize the separation between samples of different groups. Finally, weights can be extracted and a correlation with each variable computed. This is called Variable Importance in Projection and the higher the value the more the variable will have contributed to the classifier model. We filtered variables deemed significant for the classification task using a VIP threshold value of 1(15) for each one of the models. The lists were then used to perform compound set enrichment analysis.

To analyse the list produced for each group we first generated a Reactome database using files freely available from the Reactome website (https://reactome.org/download-data, lowest level pathway files). As our variables were of different type (transcriptomics, proteomics and metabolomics) we generated a merged pathway list using different Reactome files. Pathways were then selected for analysis if they had 10 or more compounds and no more than 500 as they were deemed neither robust nor informative. We then used Fisher's exact test to obtain a p-value per pathway based on the number of matches present in our list and the total

number of compounds considered initially to be included in the list of interest. For each list p-values obtained from this test were then corrected using an FDR based strategy and represented applying a threshold of .001. Following the same strategy, time points 24 and 48 were also analysed and results added alongside the ones obtained for time point 0.

Validation in an independent dataset

PLS-DA models with 3 components were created for each one of the IMOFAP group but using solely metabolites that were available for the KAPVAL cohort as well. VIP scores were computed and a threshold of 25 variables was set as a maximum of features to be included in each model. An optimal number of variables between 3 and 25 was selected based on maximum accuracy values.

To determine allocations for the KAPVAL cohort we applied each one of the PLS-DA models to all KAPVAL individual and allocated them to the group from which they were the closest to.

To compare the biology between the groups of the two cohorts we computed average values per variable per group and compared them between the cohorts as inspired by Sweeney and al.(16). We only compared metabolites that were not included in the PLS-DA models used to classify KAPVAL individuals (369 metabolites). To perform the comparison between average values we computed Spearman's correlation (comparison of ranks, p-values computed using a t-distribution) for each pairwise comparison of groups from both cohorts. Z-score values for the IMOFAP cohort and for KAPVAL allocated samples can be visualised at http://baillielab.net/pancreatitis/validation (username: pancreas and password: review).

Comparison with an external dataset

To compare obtained endotypes with external data, ARDS endotypes definitions(17) were selected. Matched variables between the IMOFAP and chosen ARDS datasets were compared using variable ranks and Spearman's correlation coefficients.

When computing Spearman's correlation coefficients and associated p-values for, we used the statsmodels and scipy modules in Python.

Supplementary Results

Clustering results

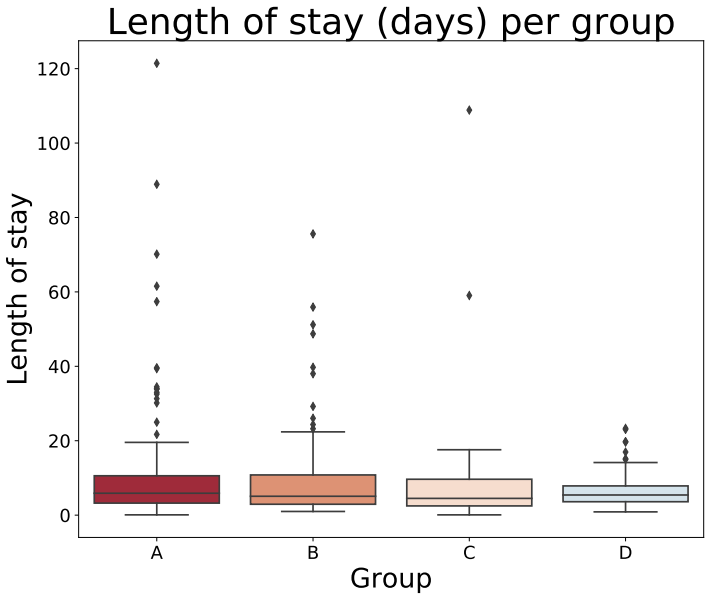
For the AUC+PCA method, the total percentage of variance explained by the first two components used was 51.5%, comprising 40.2% for PC1 and 11.3% for PC2. For this four-group solution, the silhouette score(18) based on the distance matrix was 0.39 (range 0.23 to 0.57).

Moreover, we generated a clustering for the AUC combined with PCA method using the residuals from a linear model which included gender, age and time of onset. The 4-cluster solution obtained using the corrected data was compared to the chosen partition and showed a high level of similarity (JI = 0.82 and distance matrices correlation, using a Mantel test, showed a correlation of 0.91 with an associated p-value of .01).

Based on PLS-DA results, for all compounds with a VIP associated value greater or equal to 2 averaged time profiles and AUC values can be obtained via an online page that can be accessed through the following address: http://baillielab.net/pancreatitis/, username: pancreas and password: review) along with some clinical and cytokine measures (filtered variables based on ANOVA results and a threshold of .05).

It is interesting to note that correlation between distance matrices obtained using only one type of omics and the complete distance matrix (as used to highlight the presented clusters), using Mantel test, suggest that most of the signal can be attributed to transcriptomics data (correlation value equal to 1, FDR-corrected p-value <.001). Moreover, the correlation coefficient between the metabolomics-based distance matrix and the complete distance matrix shows a value of 0.22 (FDR-corrected p-value 0.03), suggesting some contribution.

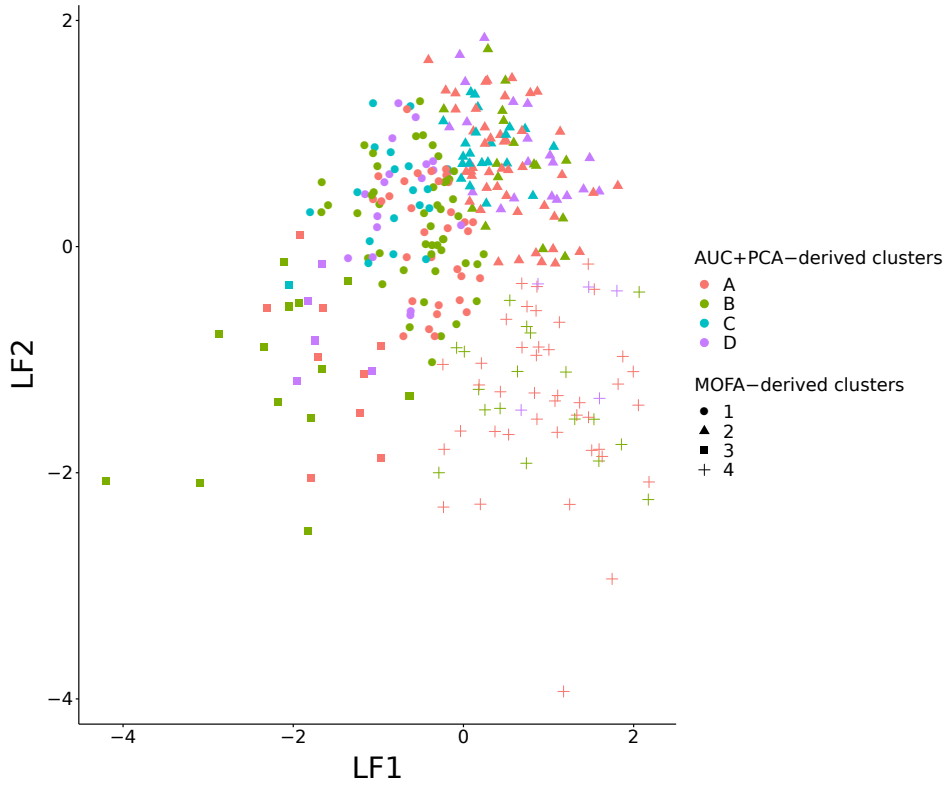
Length of stay varied between the groups with the reported values of endotype A globally higher when compared to others (median values were 5.9, 5.1, 4.5 and 5.3 days respectively, for each group with corresponding Q1-Q3 3.2-12.8, 2.8-10.8, 2.5-9.7 and 3.3-7.8, in days **Supplementary Figure B**).



Supplementary figure B. Length of stay for the KAPVAL cohort. Box plot of length of hospital stay (in days) per identified endotype for KAPVAL allocated samples. The bars represent 95% confidence intervals and the central bar depicts the median value. The x axis shows endotypes to which KAPVAL samples were allocated (A, B, C and D) and the y axis, the length of stay for each one of the allocated samples, in days.

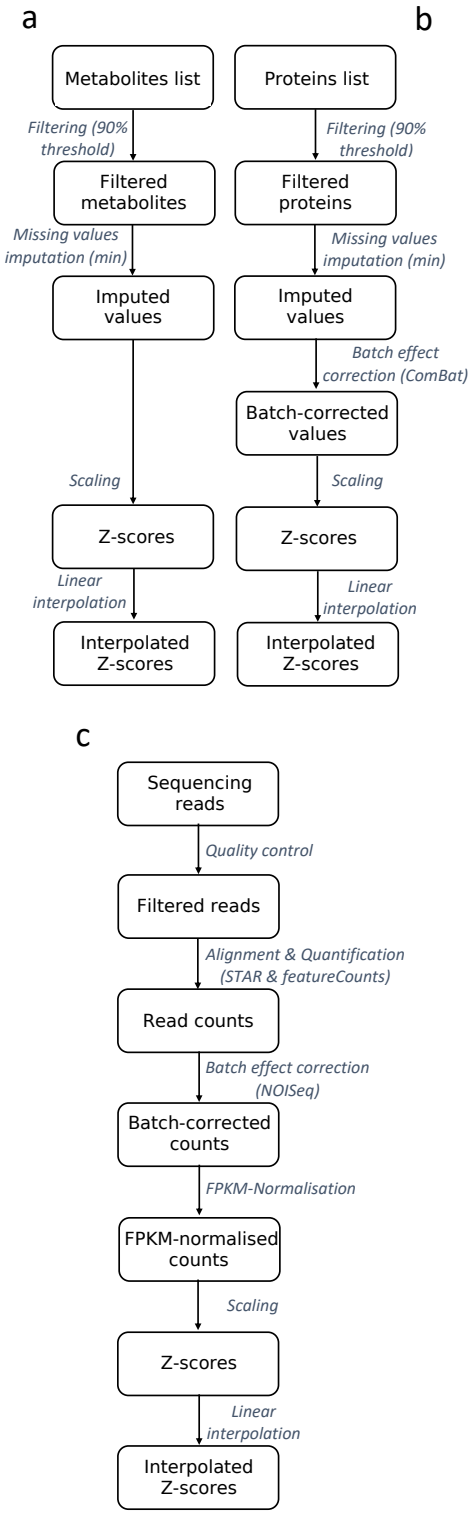
External validity

We also noticed a great difference in structure between our endotypes and the ones highlighted using MOFA when applied to baseline KAPVAL data (Jaccard index 0.22, **Supplementary Figure C**), confirming that baseline data alone cannot be used directly to highlight the structure of AP endotypes and that previous knowledge of patient trajectories is required to generate the endotype models. However, once established the models could be applied to presentation data without the further need to collect several time points.



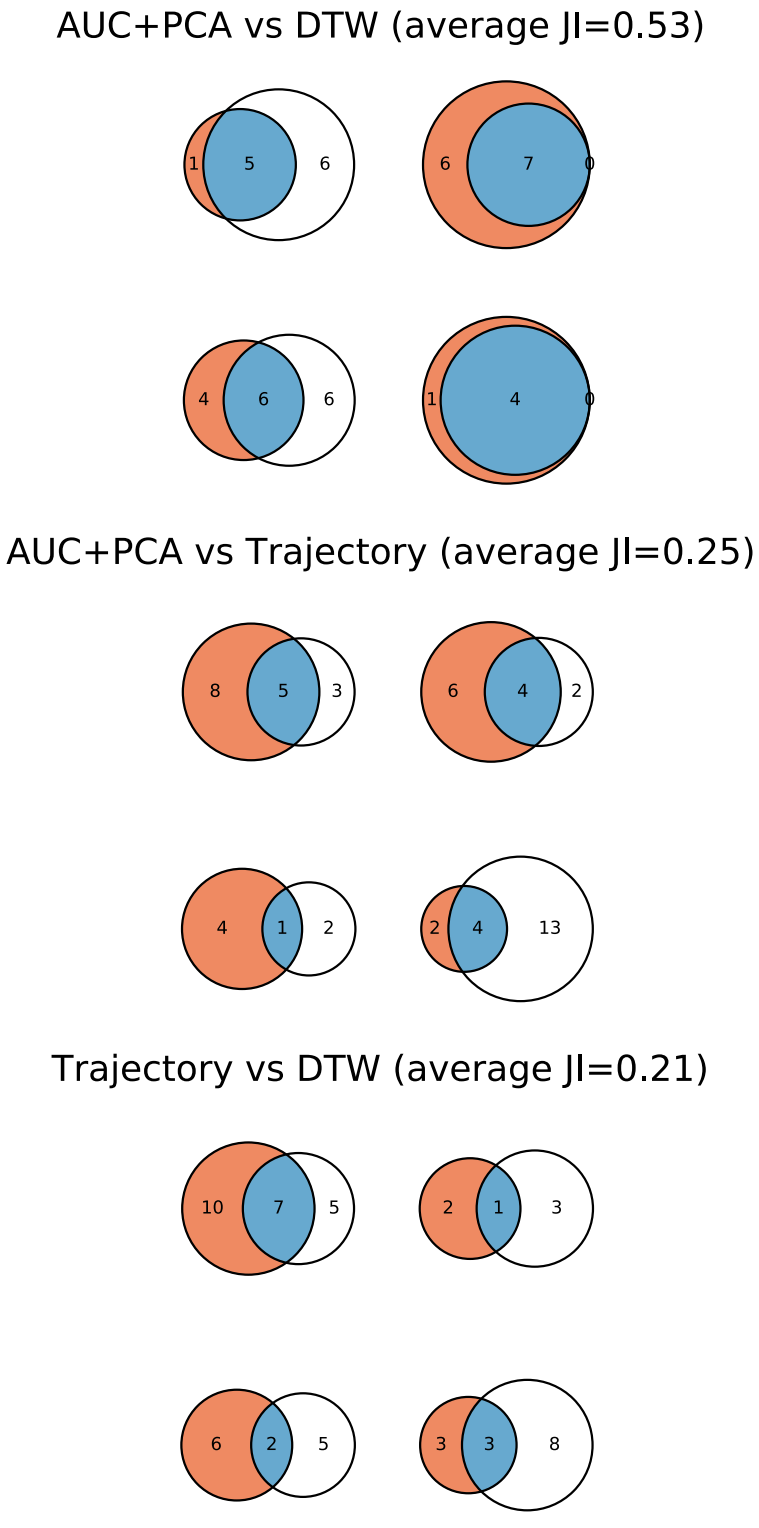
Supplementary figure C. MOFA tools results compared to highlighted clusters with KAPVAL cohort data. Using KAPVAL metabolite data and selecting a 4-cluster solution, comparison of results. KAPVAL samples were allocated to the four endotypes identified (using AUC combined with PCA) and as illustrated in Figure 3a. Colours indicate the results of the allocation obtained using PLS-DA models and shapes show MOFAtools results when choosing four clusters and two latent factors. Values of KAPVAL samples on the latent factors are plotted on the figure.

Supplementary Figures



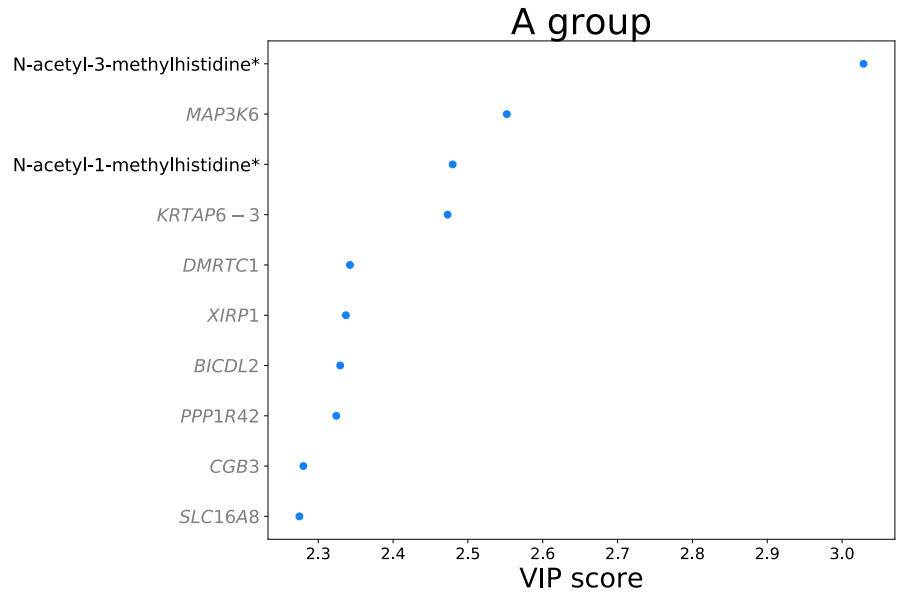
Supplementary Figure 1. Pre-processing summary. For metabolomics (panel a), proteomics (panel b) and transcriptomics (panel c), an overview of the pre-processing steps presented in the Materials and Methods section is shown. Metabolomics and proteomics data were pre-

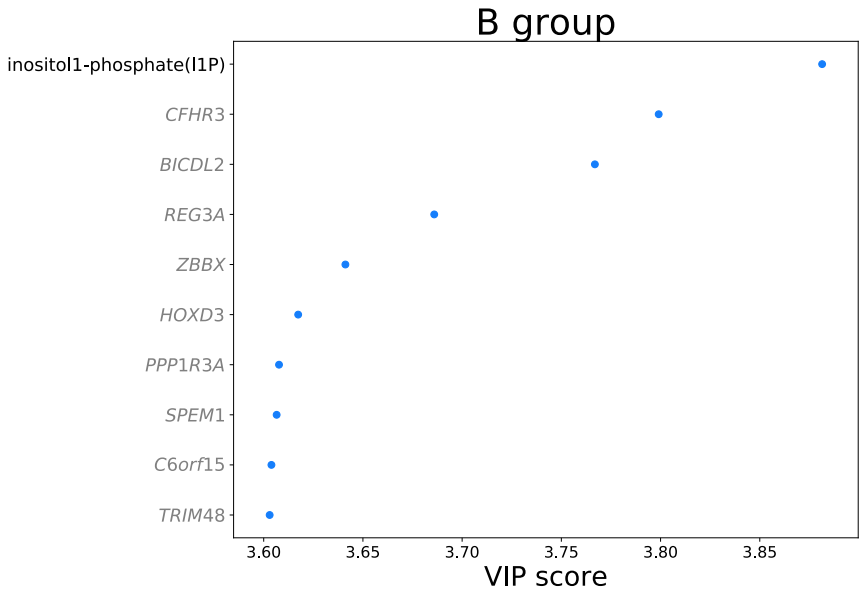
processed in a very similar way. A filtering was applied to discard variables with 90% or more missing values. This was followed by an imputation step using variables minimum value, as a missing value indicates a value below the detection limit. To remove the batch effect present in the proteomics data, ComBat was used. For both metabolomics and proteomics, a Z-score scaling was applied and was followed by linear interpolation. Quality control was done for raw transcriptomics data (using FASTQC results to trim reads with low ends using a phred threshold of 20 and discard reads shorter than 25 bp). Alignment and quantification using respectively STAR and featureCounts was performed and was followed by FPKM-normalisation. As with metabolomics and proteomics data, Z-score scaling and linear interpolation were applied.

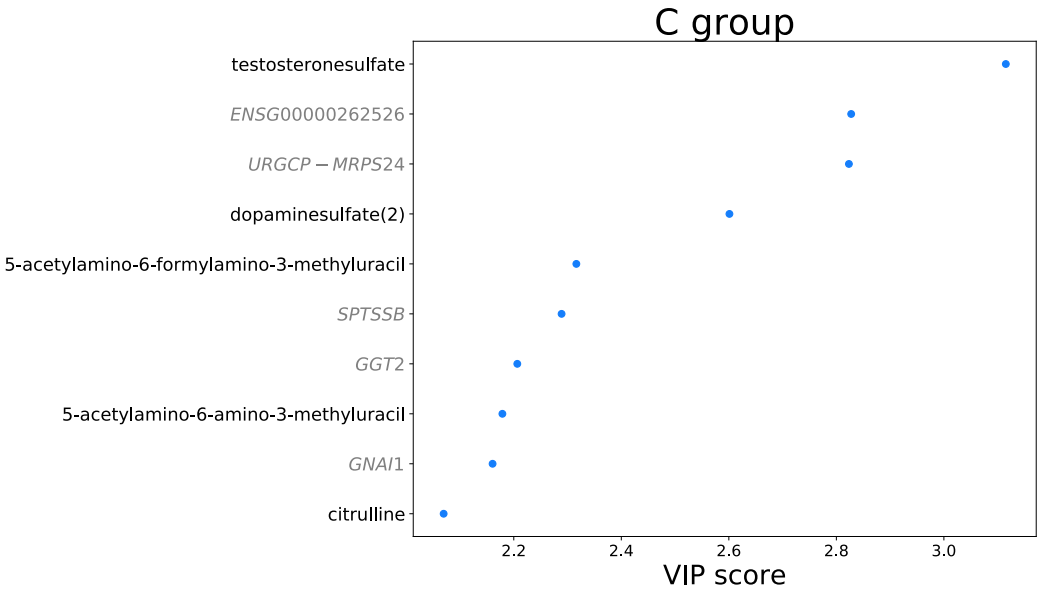


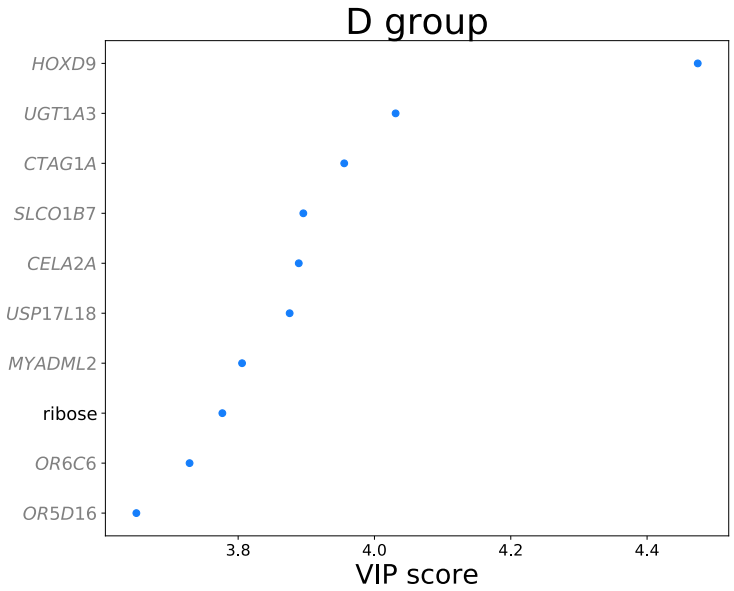
Supplementary figure 2. Overlap between the different clustering solutions. For the optimum clustering (AUC+PCA-based, consisting of 4 clusters), comparison with other four-

group solutions obtained using the two other main methods (Trajectory and DTW). Average Jaccard Index (JI) values (one value is computed for each of the four groups) are reported for each comparison to assess the overlap. Numbers within the circles correspond to numbers of individuals. Group matching (to determine which group of the first method should be compared to a group of the second method) between two methods was performed by trying all possible pairwise combinations and choosing the one producing the highest average Jaccard index. For any A vs B comparison, blue sections correspond to the number of cluster elements in common between solutions A and B. Orange and white sections refer to cluster elements not in common but respectively part of solution A and B. DTW refers to dynamic time warping, AUC+PCA to area-under-the-curve combined to principal component analysis and trajectory to trajectory in principal component analysis space.

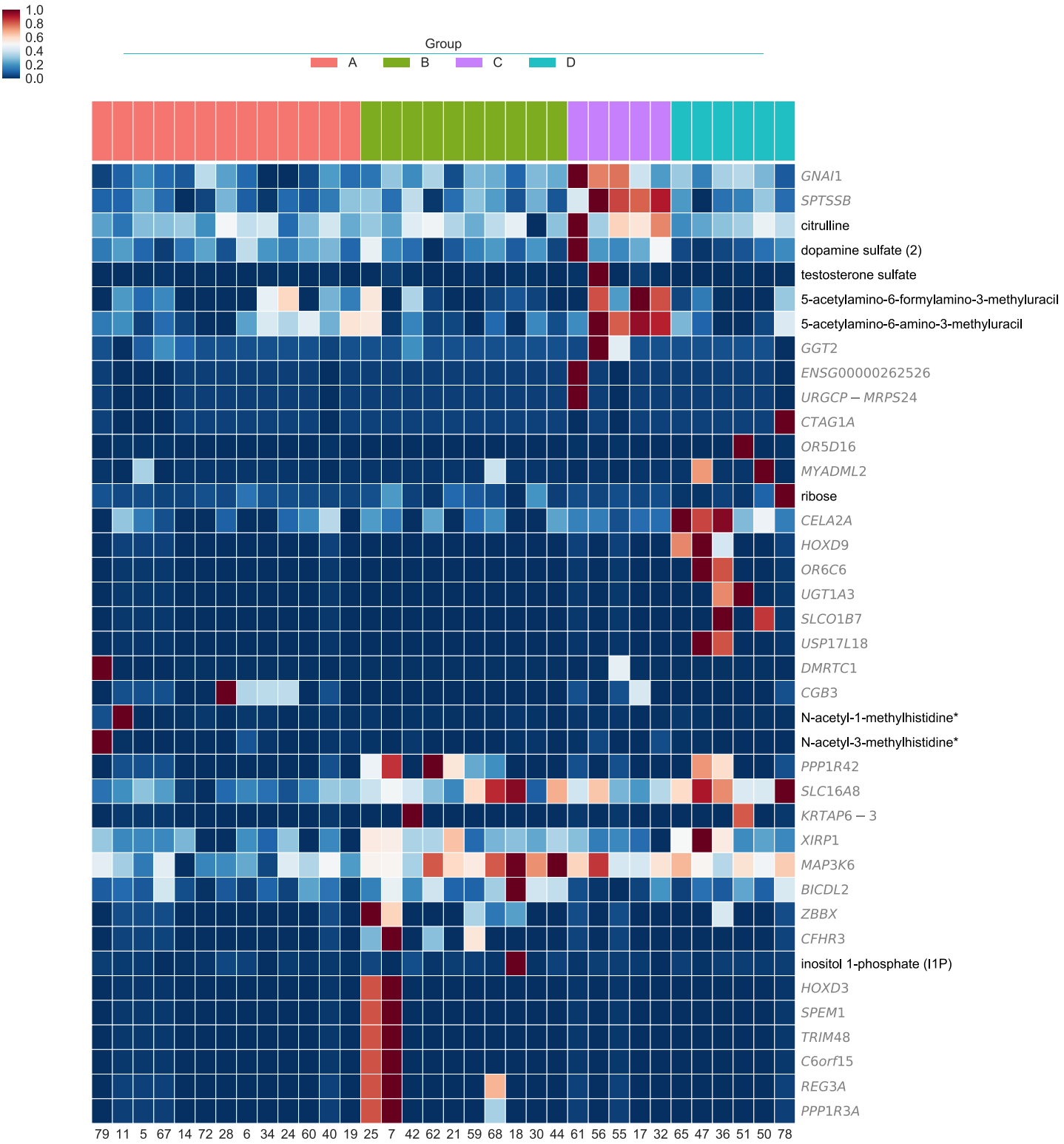

A group
N-acetyl-3-methylhistidine*
MAP3K6
N-acetyl-1-methylhistidine*
KRTAP6 − 3
DMRTC1
XIRP1
BICDL2
PPP1R42
CGB3
SLC16A8
2.3
2.4
2.5
2.6
2.7
2.8
2.9
3.0
VIP score


B group
inositol1-phosphate(I1P)
CFHR3
BICDL2
REG3A
ZBBX
HOXD3
PPP1R3A
SPEM1
C6orf15
TRIM48
3.60
3.65
3.70
3.75
3.80
3.85
VIP score


C group
testosteronesulfate
ENSG00000262526
URGCP − MRPS24
dopaminesulfate(2)
5-acetylamino-6-formylamino-3-methyluracil
SPTSSB
GGT2
5-acetylamino-6-amino-3-methyluracil
GNAI1
citrulline
2.2
2.4
2.6
2.8
3.0
VIP score

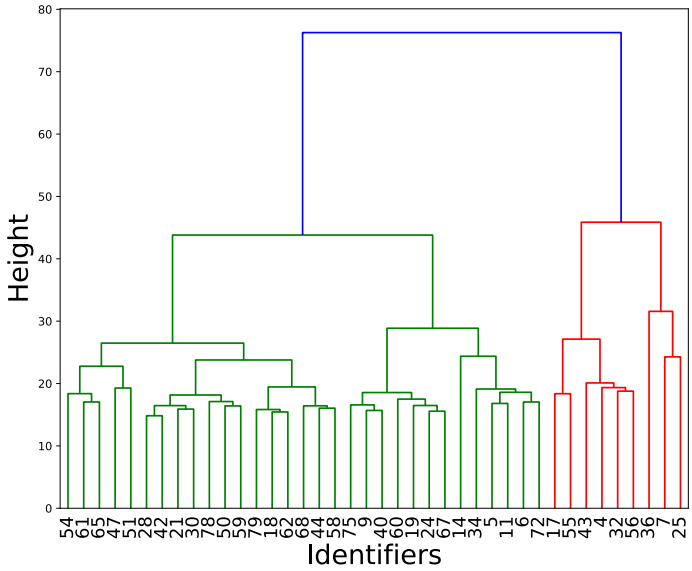


Supplementary figure 3. PLS-DA top variables for each identified endotype. The top 10 variables from the four PLS-DA models (for clusters A, B, C and D) are presented. Top variables were extracted based on VIP (Variable Importance in Projection) values**.** Names on the y axis refer to gene (in grey italic) or metabolite compounds. Values reported on the x axis are VIP values and are computed using correlation values between PLS-DA components and group labels. A high VIP value represents a high contribution of the variable to the model.

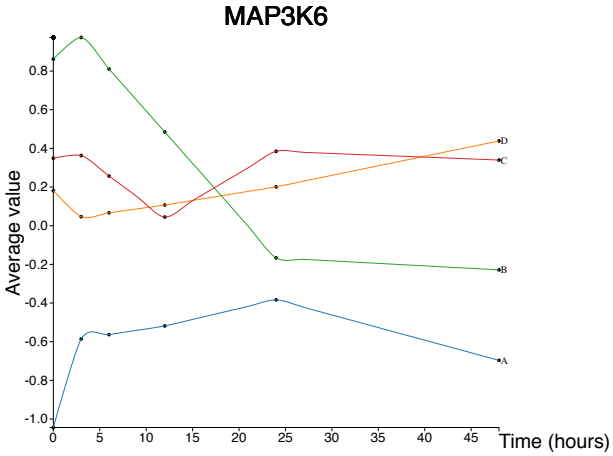


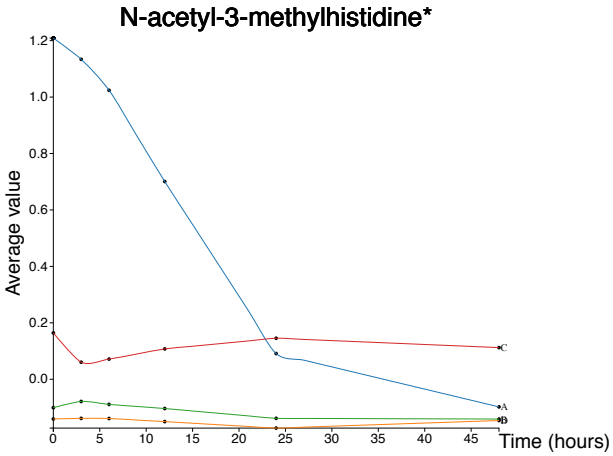
Supplementary figure 4. AP endotypes overview for VIP-selected variables. Heatmap of normalised and scaled variable values of the top 10 variables (as reported in Supplementary figure 3), for each identified group. Compound names, on the y axis, are reported in grey italic when corresponding to genes, all others refer to metabolites. Patient identifiers are reported on the bottom x axis and colours represent endotype allocations on the top x axis. For visualisation

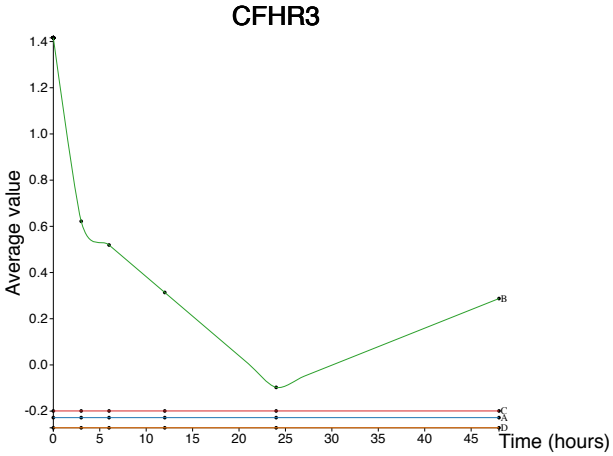
purposes row values were scaled between 0 and 1. Colours are representative of the range of values observed. A star indicates a patient who did not survive.

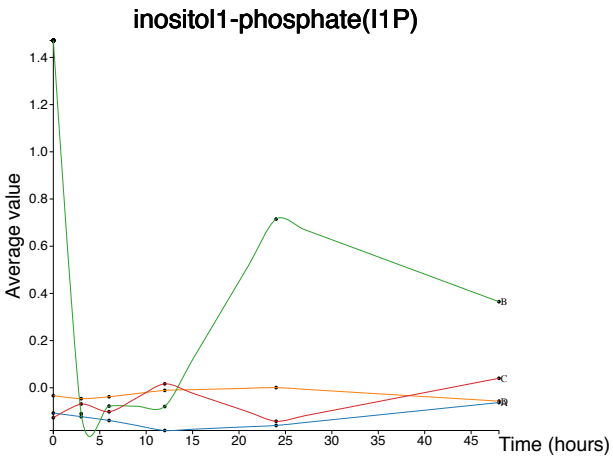


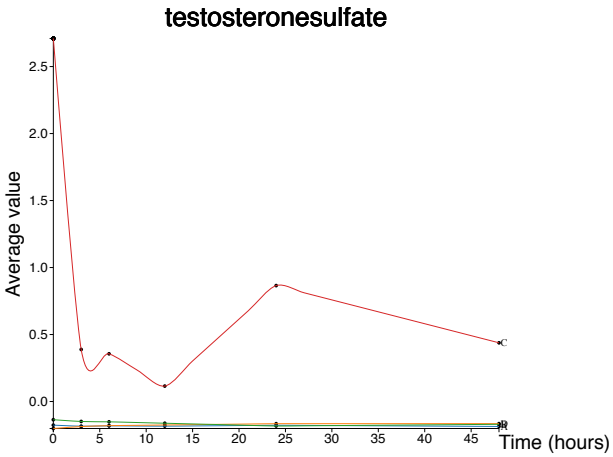
Supplementary figure 5. Hierarchical clustering result for time point 0 data. Using only time 0 data and choosing two as an arbitrary number of clusters, dendrogram based on Euclidean distances and Ward's algorithm. Patient identifiers are reported on the x axis and distances on the y axis.

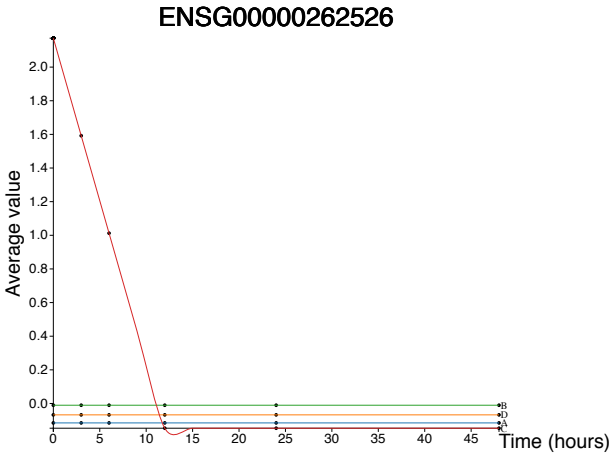

MAP3K6
Average value
Time (hours)
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
0
5
10
15
20
25
30
35
40
45
D
C
B
A

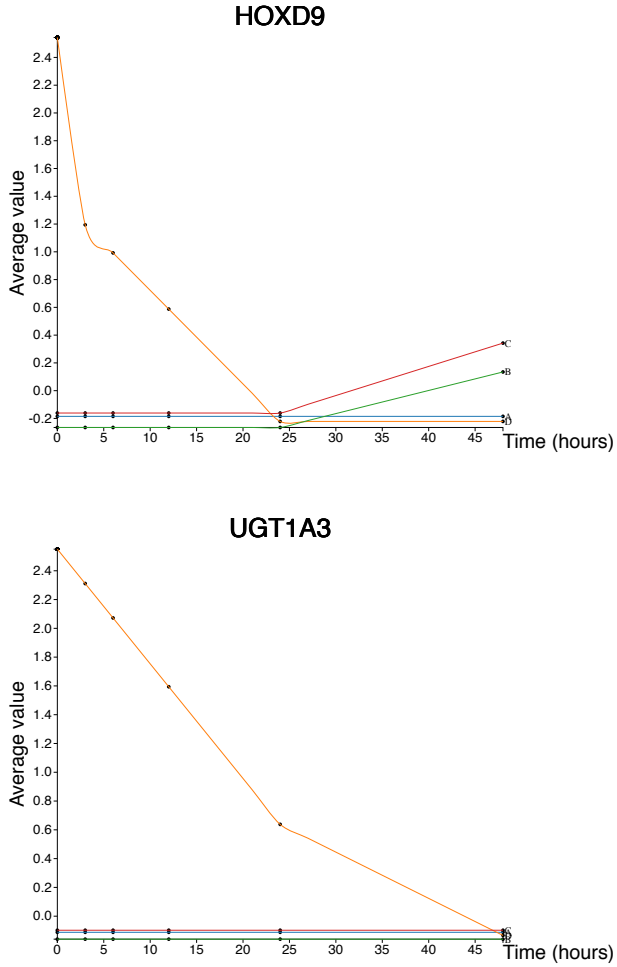

N-acetyl-3-methylhistidine*
Average value
Time (hours)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
35
40
45
C
A
B


CFHR3
Average value
Time (hours)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
5
10
15
20
25
30
35
40
45
B
C
A
D

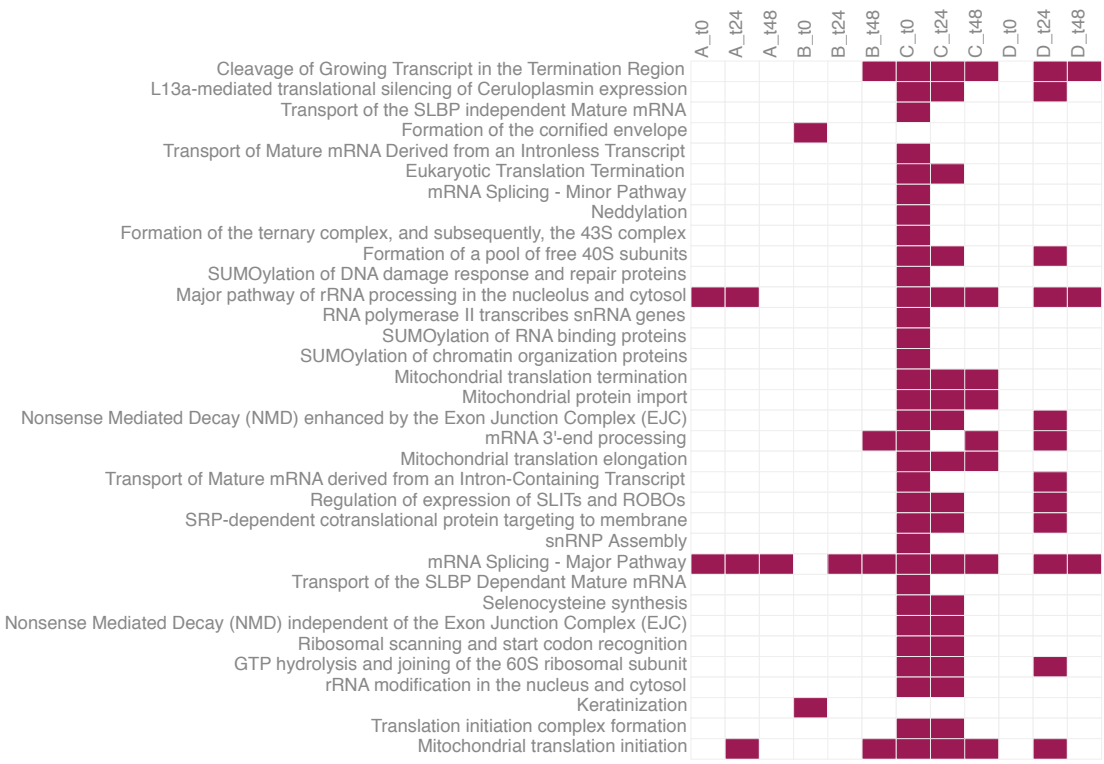

inositol1-phosphate(I1P)
Average value
Time (hours)


testosteronesulfate
Average value
Time (hours)

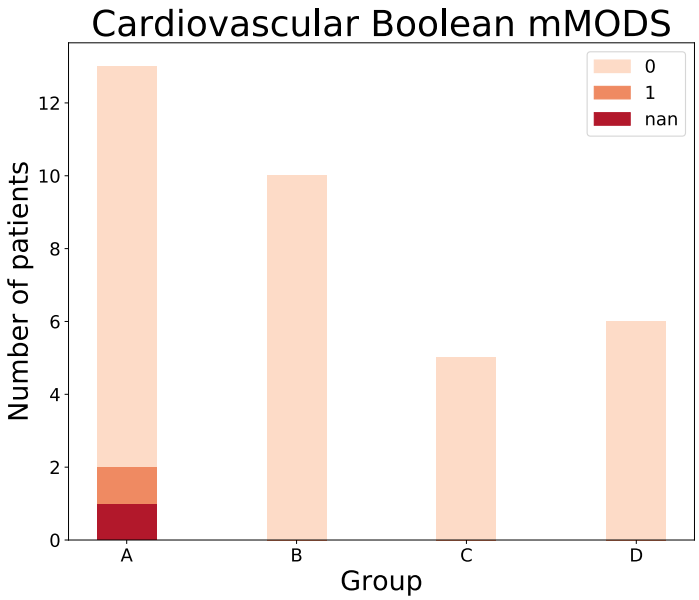

ENSG00000262526
Average value
Time (hours)

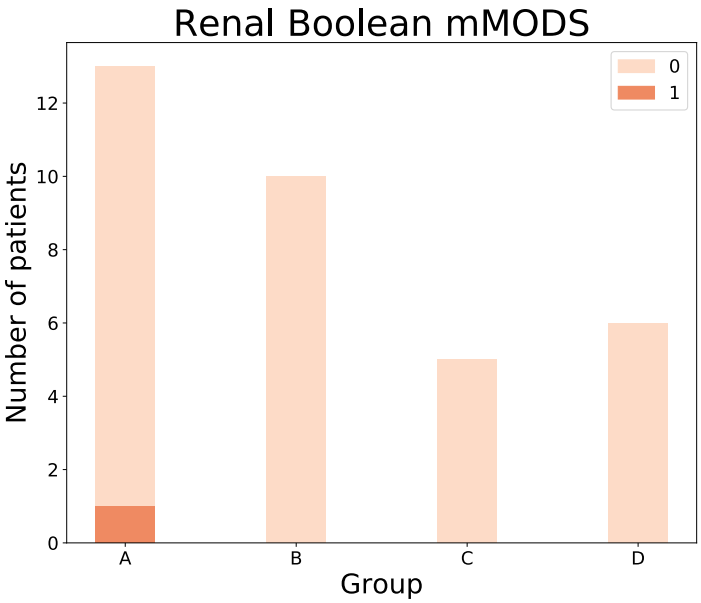


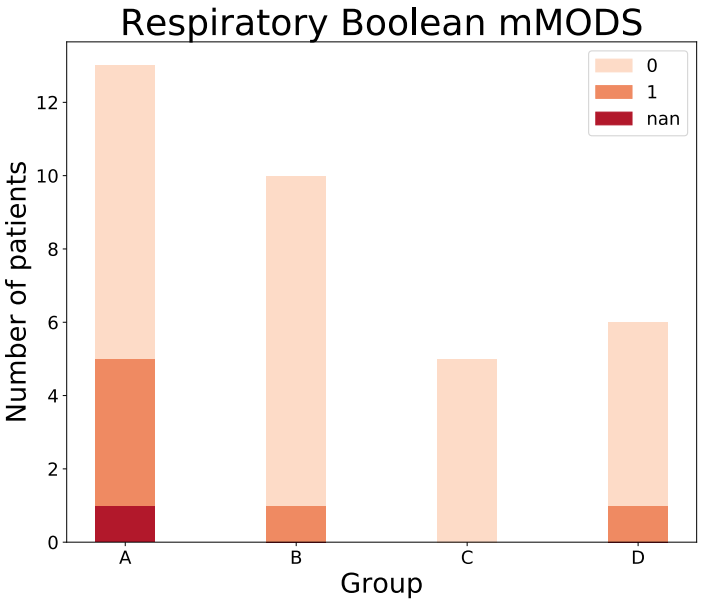
Supplementary figure 6. Time profiles per group for selected variables. Time series generated from average Z-score value per time point per group identified using our AUC-PCA strategy. Imputed values (using linear imputation and 3-hour intervals) were used to generate the plots. The top 2 variables per group were selected from PLS-DA results using VIP scores and consisted of 5 genes and 3 metabolites. The average Z-score values per time point are represented using the y axis. Points represent sampling times. The x axis shows time in hours and curves were obtained using Cardinal splines. Graphs were produced with http://baillielab.net/pancreatitis/ (username: pancreas and password: review, graphs generated by clicking on variable names) using D3.js.



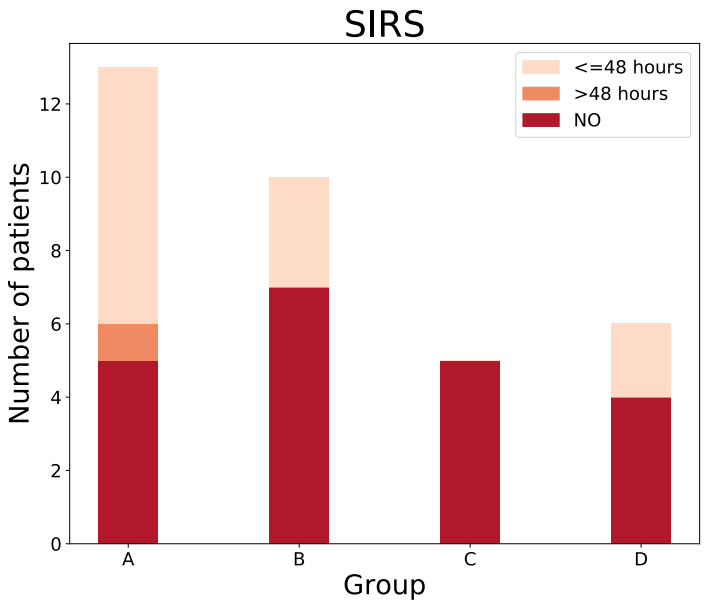
Supplementary figure 7. Enrichment results. Significant pathway terms (filtered using an adjusted p-value threshold of 0.01%) from enrichment results for each identified group based on variables lists selected using VIP scores. For the enrichment analysis we selected variables with a VIP score value equal to or greater than 1, based on PLS-DA models. Pathway data was fetched from the Reactome pathway database. Time point 0 was selected for assessment and all data types were included (metabolites, proteins and genes). Results for time points 24 and 48 hours are reported as well. Fisher's test was used to generate p-values. Pathway full names are reported in the figure. Columns represent results for one endotype at a selected time point. A coloured cell represents a significant value and a white cell a non-significant value (based on FDR-adjusted p-values).


Cardiovascular Boolean mMODS
0
1
nan
Number of patients
0
2
4
6
8
10
12
A
B
C
D
Group

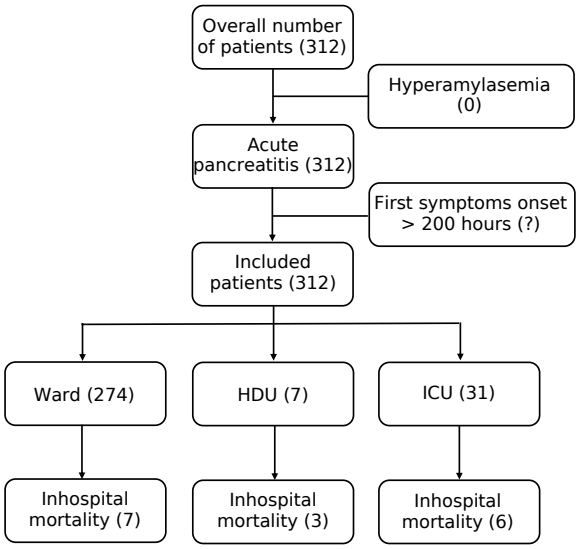

Renal Boolean mMODS
0
1
Number of patients
0
2
4
6
8
10
12
A
B
C
D
Group


Respiratory Boolean mMODS
0
1
nan
Number of patients
0
2
4
6
8
10
12
A
B
C
D
Group

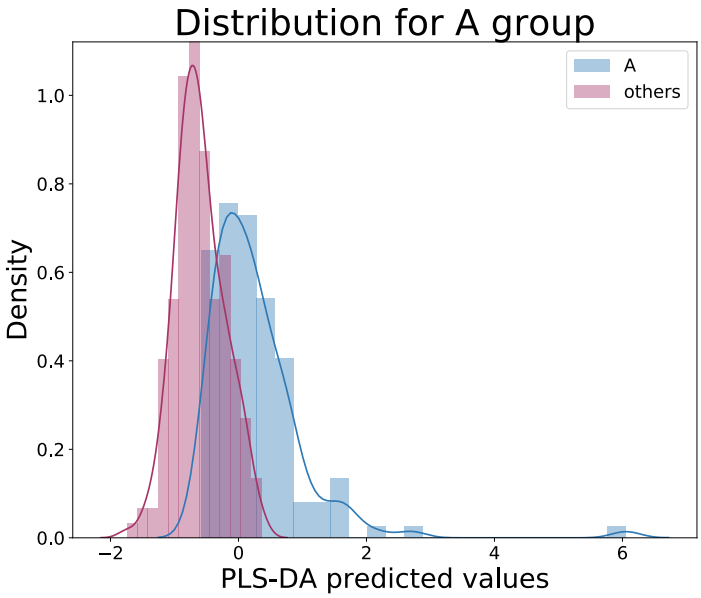
Supplementary figure 8. Modified MODS distribution per endotype. For each one of the four endotypes, the number of patients in each modified MODS category (respectively cardiovascular, renal and respiratory) is shown using a colour code. Patients were categorised in two classes, as described in the figure's legend. A value of 1 was given to any patient having a modified MODS greater than one at time point 0. A value of 0 was given otherwise. *nan* represents a missing value. The x axis was used to represent the identified endotypes and the y axis to show the number of patients, per endotype, falling in each one of the described modified MODS categories.

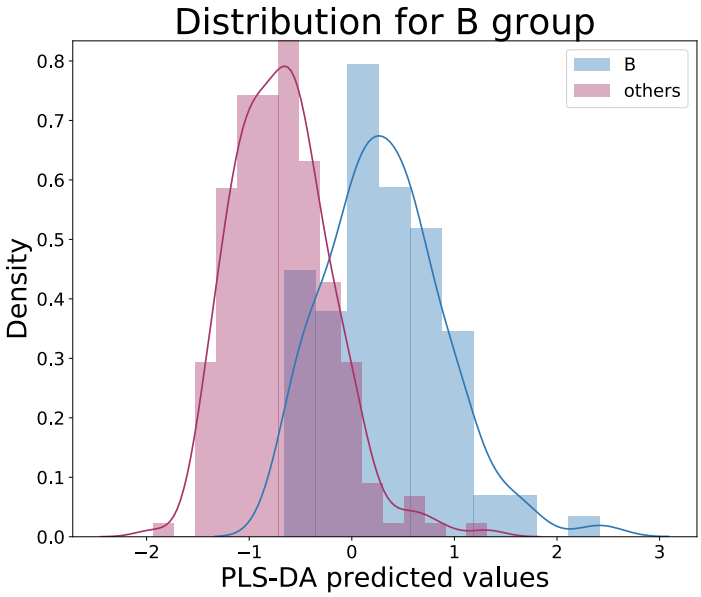


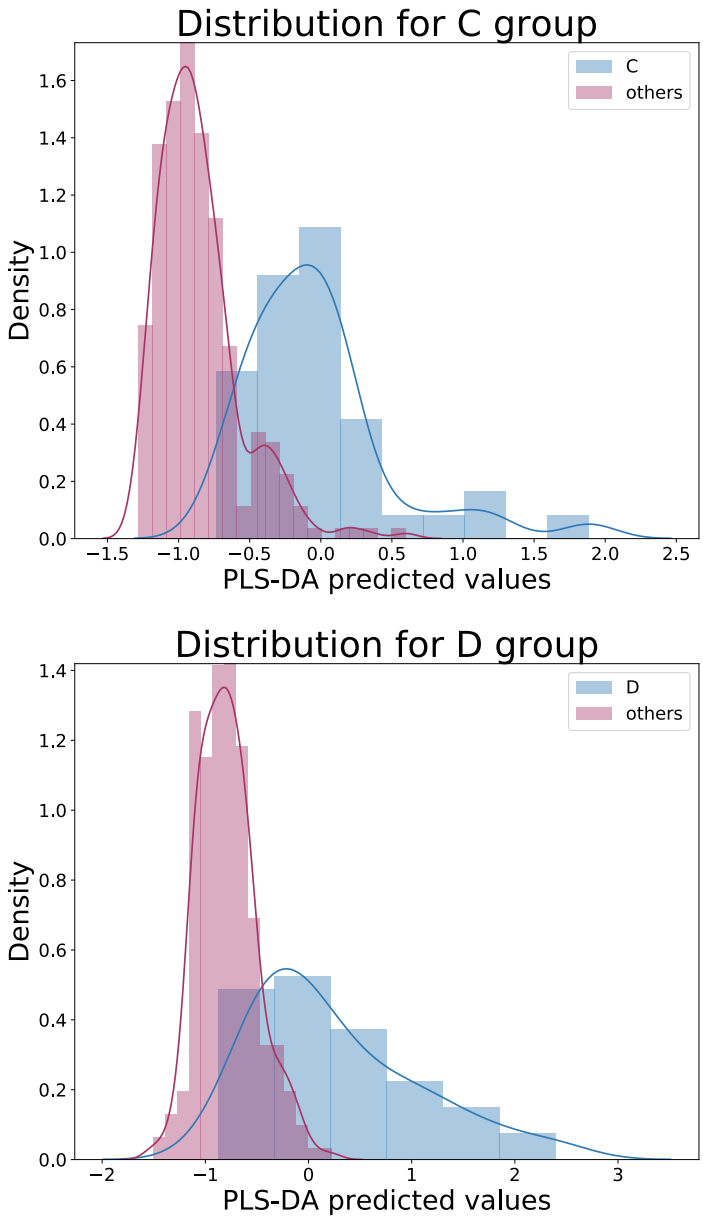
Supplementary figure 9. SIRS distribution per endotype. For each one of the four endotypes, the number of patients in each SIRS (Systemic Inflammatory Response Syndrome) category is shown using a colour code. Patients were categorised in three classes, as described in the figure's legend, corresponding respectively to SIRS persisting for 48 or less consecutive hours, SIRS persisting for more than 48 hours and no reported SIRS. The x axis was used to represent the identified endotypes and the y axis to show the number of patients, per endotype, falling in each one of the described SIRS categories.



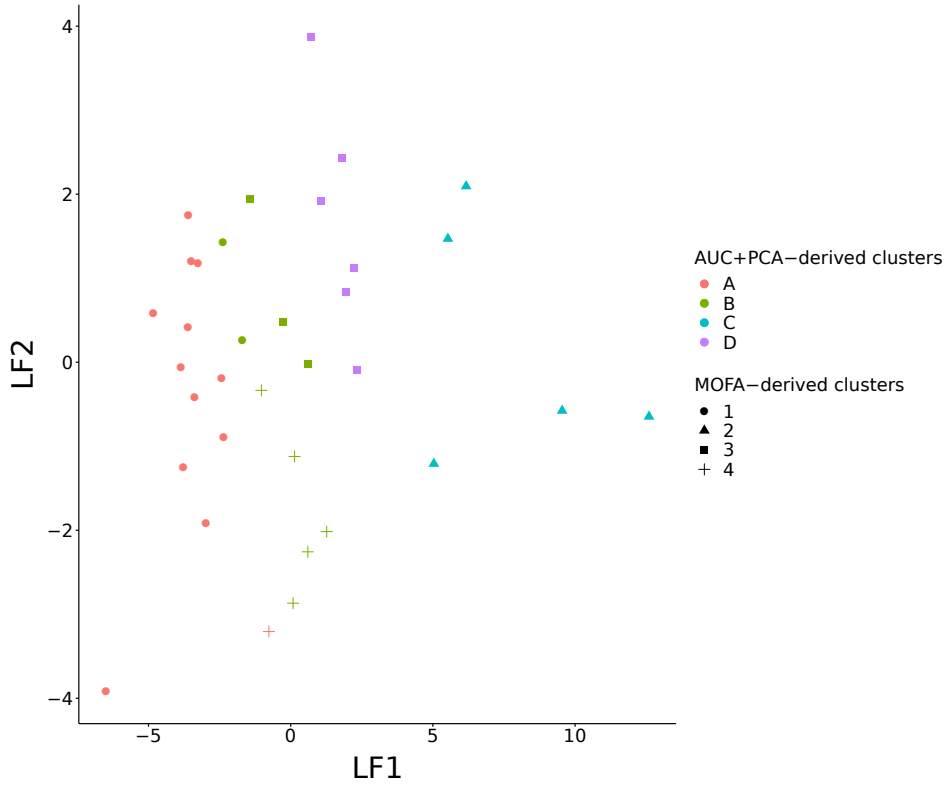
Supplementary figure 10. Flow diagram for the KAPVAL cohort. Study flow chart for patient samples and data from the KAPVAL study included in the analysis showing cohort structure. Individuals were-prefiltered for hyperamylasemia and thus all 312 individuals were kept for the analysis. As opposed to the IMOFAP cohort, symptoms onset time was not recorded and thus no filtering could be performed in that regard.


Distribution for A group
A
others
Density
PLS-DA predicted values


Distribution for B group
B
others
Density
PLS-DA predicted values



Supplementary figure 11. PLS-DA predicted values for each identified endotype. PLS-DA predicted values distribution plots for all four identified endotype. When building the models, a value of 1 is assigned to the current endotype and a value of -1 to all others. For each endotype, distribution of predicted values, as returned by the PLS-DA procedure, for assigned (represented in blue, assignment to the best matching group) and unassigned KAPVAL individuals (depicted in pink) are represented. Kernel density estimates are fitted for both unassigned and assigned samples in each endotype graphs.



Supplementary figure 12. MOFA tools results compared to highlighted clusters with IMOFAP cohort data. Comparison of MOFAtools results with AUC+PCA-derived clusters. AUC values were used as input (to be consistent with the AUC+PCA strategy) and a 4-cluster solution was extracted from MOFA results using the first two latent features (LF1 and LF2). Colours are representative of clusters described in this paper and shapes of MOFAtools predicted allocations. LF1 and LF2 axes represent latent factors as defined by the MOFA algorithm. Values for the two latent factors are plotted.

Supplementary Tables

Supplementary Table 1. Demographics. Summary clinical data for clustered participants (n=34) of the IMOFAP cohort. The cohort is fully described in Skouras et al (for n=57 AP patients, including n=3 patients with a symptoms onset at recruitment equal to or greater than 200 hours and 20 individuals with data available only for one time point).

Number of patients		**34**
Gender		
	Male	55.88% (n=19)
Age (years)		
	Median	58.30
	IQR	23.05 (49.60-72.65)
BMI		
	Median	27
	IQR	7.75 (23-30.75)
Source of recruitment		
	A&E	94.12% (n=32)
	Other	5.88% (n=2)
Length of hospital stay (days)		
	Median	5
	IQR	3 (4-7)
Aetiology		
	Gallstones	50.00% (n=17)
	Alcohol	29.41% (n=10)
	Other	20.59% (n=7)
Charlson index (time point 0)		
	Median	2
	IQR	3 (1-4)
Inhospital mortality (binary)		
	1	5.88% (n=2)
Time onset recruitment (hours)		
	Median	22.23
	IQR	40.86 (11.96-52.81)
Alcohol use		
	Current	47.06% (n=16)
	Previous	8.82% (n=3)
	None	44.12% (n=15)
Smoking		
	Current	29.41% (n=10)
	Previous	14.71% (n=5)
	None	55.88% (n=19)
Critical care admission (binary)		
	1	8.82% (n=3)
APACHE II day 1		
	Median	10
	IQR	5 (8-13)
Previous AP		
	0	61.76% (n=21)
	1	29.41% (n=10)
	2	5.88% (n=2)
	3 or more	2.94% (n=1)
CRP (mg/L) (time point 0)		
	Mean	76.21
	SD	98.86

Supplementary table 2. Clustering similarities for 3-cluster solutions. For each time-series-based method (AUC combined with PCA, Trajectory within PCA-space and dynamic time warping), based on 3-cluster solutions, pairwise comparisons were made. Average Jaccard index values (averaged over the 3 clusters being compared) are reported in the following table. As described for Supplementary figure 2, for a given pairwise comparison, clusters between two solutions were matched by maximising the average Jaccard index value. Values are displayed for each pairwise comparison and give an idea of the similarity between clustering obtained with different methods. Network graphs allowing to compare the different solutions can be generated using http://baillielab.net/pancreatitis/networks (username: pancreas and password: review).

Average Jaccard index	AUC+PCA	PCA+Trajectory	Dynamic time warping
AUC+PCA	/	/	/
PCA+Trajectory	0.31	/	/
Dynamic time warping	0.63	0.27	/

Supplementary table 3. Clustering similarities for 5-cluster solutions. For each time-series method (AUC combined with PCA, Trajectory within PCA-space and dynamic time warping), based on 5-cluster solutions, comparisons were run using Jaccard index values (computed as described in Supplementary table 2). Reported values are averaged Jaccard indexes. Values are displayed for each pairwise comparison and gives an idea of overlap between the different clustering solutions. Network graphs are available at http://baillielab.net/pancreatitis/networks (username: pancreas and password: review).

Average Jaccard index	AUC+PCA	PCA+Trajectory	Dynamic time warping
AUC+PCA	/	/	/
PCA+Trajectory	0.24	/	/
Dynamic time warping	0.46	0.21	/

Supplementary table 4. Top compound sets. Using glm in R, top 20 pathways (using KEGG data for gene, protein and metabolite data and FANTOM5 data for gene and protein data) for clusters obtained using the AUC combined with PCA method. FDR-corrected p-values obtained are reported along with pathway identifiers and complete names. Input data corresponds to time point 0.

Pathway	FDR-corrected p-value
hsa00190 Oxidative phosphorylation	<.001
hsa00230 Purine metabolism	<.001
hsa00240 Pyrimidine metabolism	<.001
hsa00510 N-Glycan biosynthesis	<.001
hsa00970 Aminoacyl-tRNA biosynthesis	<.001
hsa03008 Ribosome biogenesis in eukaryotes	<.001
hsa03010 Ribosome	<.001
hsa03013 RNA transport	<.001
hsa03015 mRNA surveillance pathway	<.001
hsa03018 RNA degradation	<.001
hsa03040 Spliceosome	<.001
hsa04010 MAPK signaling pathway	<.001
hsa04110 Cell cycle	<.001
hsa04120 Ubiquitin mediated proteolysis	<.001
hsa04141 Protein processing in endoplasmic reticulum	<.001
hsa04142 Lysosome	<.001
hsa04144 Endocytosis	<.001
hsa04146 Peroxisome	<.001
hsa04660 T cell receptor signaling pathway	<.001
hsa00280 Valine, leucine and isoleucine degradation	<.001

Supplementary table 5. Heatmap compounds characteristics. Compounds detailed table for heatmap presented in figure 2a. As ordered in figure 2a. Complete gene names were fetched using the GeneCards resource and additional information using online resources such as HUGO Gene Nomenclature Committee, NCBI EntrezGene, UniProtKB, Ensembl, NCBI PubChem, NCBI PubMed and Google Search. Compounds reported in grey italic correspond to genes and others to metabolites.

Heatmap compound	Complete name	Additional information
GNAI1	G Protein Subunit Alpha I1	N-acetyl transferase activity
SPTSSB	Serine Palmitoyltransferase Small Subunit B	Tricarboxylic acid cycle(19)
Citrulline	/	Sphingolipid biosynthesis
Dopamine sulfate (2)	/	Gastrointestinal dopamine metabolism(20)
Testosterone sulfate	/	
5-acetylamino-6-amino-3-methyluracil	/	Caffeine metabolism(21)
5-acetylamino-6-formylamino-3-methyluracil	/	
GGT2	Gamma-Glutamyltransferase 2	γ-glutamyl transferase; Glutathione homeostasis
URGCP-MRPS24	URGCP-MRPS24 Readthrough	
ENSG00000262526	/	Protein coding
OR5D16	Olfactory Receptor Family 5 Subfamily D Member 16	
CTAG1A	Cancer/Testis Antigen 1A	
MYADML2	Myeloid Associated Differentiation Marker Like 2	
Ribose	/	
CELA2A	Chymotrypsin Like Elastase Family Member 2A	Pancreatic elastase-2
HOXD9	Homeobox D9	
OR6C6	Olfactory Receptor Family 6 Subfamily C Member 6	
UGT1A3	UDP Glucuronosyltransferase Family 1 Member A3	Associated with Gilbert-type hyperbilirubinemia(22, 23)
SLCO1B7	Solute Carrier Organic Anion Transporter Family Member 1B7 (Putative)	Cysteine-type endopeptidase
USP17L18	Ubiquitin Specific Peptidase 17-Like Family Member 18	Liver-specific organic anion transporter; Bile secretion
DMRTC1	DMRT Like Family C1	
CGB3	Chorionic Gonadotropin Subunit Beta 3	
N-acetyl-1-methylhistidine*	/	Amino acid metabolism; Rhabdomyolysis; Renal failure(24)
N-acetyl-3-methylhistidine*	/	
PPP1R42	Protein Phosphatase 1 Regulatory Subunit 42	

SLC16A8	Solute Carrier Family 16 Member 8	Lactate transporter; Ketone body transporter
KRTAP6-3	Keratin Associated Protein 6-3	
XIRP1	Xin Actin Binding Repeat Containing 1	Muscle-specific actin binding protein upregulated during muscle injury(25)
MAP3K6	Mitogen-Activated Protein Kinase Kinase Kinase 6	Apoptosis signaling(26)
BICDL2	BICD Family Like Cargo Adaptor 2	
ZBBX	Zinc Finger B-Box Domain Containing	
CFHR3	Complement Factor H Related 3	Heparin-binding; Complement regulation
Inositol 1-phosphate (I1P)	/	Inositol biosynthesis
HOXD3	Homeobox D3	Increases immune cell adherence; Overexpression upregulates glycoprotein IIb/IIIa(27)
SPEM1	Spermatid Maturation 1	
C6orf15	Chromosome 6 Open Reading Frame 15	Putative heparin/fibronectin binding
TRIM48	Tripartite Motif Containing 48	Interferon-γ signalling (oxidative stress/apoptosis signal-reducting kinase 1)(28)
REG3A	Regenerating Family Member 3 Alpha	Bactericidal C-type lectin; Known as pancreatitis-associated protein(29)
PPP1R3A	Protein Phosphatase 1 Regulatory Subunit 3A	Genetic association with type 2 DM and familial partial lipodystrophy 3(30)

Supplementary Table 6. Demographics. Summary clinical data for included participants of the KAPVAL cohort (n=312). For a selected number of clinical data and blood measurements, summary values for the 312 KAPVAL samples included in the analysis, are reported in the table below.

Number of patients		**312**
Gender		
	Male	46.79% (n=146)
Age (years)		
	Median	56.00
	IQR	30.25 (40.75-71.00)
Inhospital mortality (binary)		
	1	5.13% (n=16)
Critical care admission (binary)		
	1	12.18% (n=38)
CRP (mg/L)		
	Mean	47.62
	SD	79.85

Supplementary Table 7. PLS-DA models predictors. List of predictors for each one the four PLS-DA models used to allocate KAPVAL individuals to IMOFAP-derived endotypes. Metabolomics data was used to build the predictors, hence, all reported variables correspond to metabolites.

Endotype	Included predictors
Endotype A	N2,N2-dimethylguanosine N6-carbamoylthreonyladenosine Pregnenolone sulfate N-acetylserine N6-succinyladenosine 5-methylthioadenosine (MTA) Isoleucine Kynurenate 4-acetamidobutanoate Deoxycholate Xanthosine 2-hydroxypalmitate Gamma-CEHC glucuronide* 12-HETE
Endotype B	Palmitoyl ethanolamide Isovalerate 1-methylimidazoleacetate 3-hydroxy-2-ethylpropionate N-stearoyltaurine eicosapentaenoate (EPA;20:5n3) 2-hydroxypalmitate Stearate (18:0) docosadienoate (22:2n6) choline 12-HETE Docosapentaenoate (n6DPA;22:5n6) 1,3-dimethylurate palmitate(16:0)
Endotype C	Prolylglycine beta-alanine trigonelline (N-methylnicotinate) paraxanthine testosterone sulfate 4-ethylphenylsulfate oxalate (ethanedioate)
Endotype D	Dodecanedioate Laurate (12:0) lysine gamma-glutamyltyrosine methionine dimethyl sulfone 7-methylurate Caprate (10:0) orotidine guanidinoacetate glycohyocholate

Supplementary Table 8. Subgroups details. Summary clinical data for clustered participants of the IMOFAP cohort (n=34), per cluster, at recruitment. Variables correspond to the ones selected for the comparison with ARDS (ordered as in figure 4a), summary values (mean and standard deviation) for the 34 IMOFAP samples included in the analysis, at time point 0 (recruitment), stratified by group, are reported in the table below.

Number of patients		**A (n=13)**	**B (n=10)**	**C (n=5)**	**D (n=6)**
Bicarbonate (mmol/L)					
	Mean	23.85	24.10	24.40	23.50
	SD	4.39	2.23	0.89	1.64
Systolic blood pressure (mmHg)					
	Mean	136.92	133.50	152.20	122.67
	SD	26.55	22.72	21.46	14.67
Platelet count (10^9/L)					
	Mean	195.25	198.78	314.00	227.33
	SD	88.27	47.63	74.28	86.06
Albumin (g/L)					
	Mean	31.38	35.50	33.80	33.5
	SD	6.81	4.25	4.32	3.27
FiO2 (portion of 1)					
	Mean	0.26	0.21	0.22	0.22
	SD	0.11	0.01	0.02	0.02
Glucose (mg/dL)					
	Mean	8.25	5.99	5.40	7.76
	SD	5.82	0.65	1.37	3.67
BMI (kg/m^2)					
	Mean	30.31	25.90	22.4	27.83
	SD	6.58	6.23	4.45	5.56
Age (years)					
	Mean	63.41	57.84	52.86	69.28
	SD	18.19	13.67	5.14	18.51
Sodium (mmol/L)					
	Mean	140.00	137.3	138.20	136.5
	SD	4.02	2.31	2.05	3.33
White cell count (10^9/L)					
	Mean	11.24	11.32	10.40	14.42
	SD	5.76	3.93	2.35	6.12
Haematocrit (portion of 1)					
	Mean	0.38	0.39	0.36	0.39
	SD	0.05	0.06	0.07	0.06
Temperature (°C)					
	Mean	37.27	37.15	36.22	37.02
	SD	0.75	0.45	0.75	0.70
Bilirubin (umol/L)					
	Mean	30.69	27.10	7.40	39.33
	SD	20.97	17.80	2.19	35.99
Respiratory rate (/min)					
	Mean	21.92	18.00	16.80	16.67
	SD	6.40	2.05	1.79	1.75

Creatinine (umol/L)					
	Mean	99.92	69.80	64.40	70.17
	SD	47.05	17.23	7.73	15.80
Mean arterial pressure (mmHg)					
	Mean	99.67	99.00	113.40	87.17
	SD	21.10	18.09	12.66	13.38
Heart rate (/min)					
	Mean	91.00	78.70	68.00	70.00
	SD	27.31	10.56	12.41	19.31
IL 8 CXCL8 (pg/mL)					
	Mean	129.59	9.47	8.58	11.63
	SD	244.38	8.41	4.05	9.03
IL 6 (pg/mL)					
	Mean	1645.72	126.18	35.07	55.52
	SD	2704.55	168.66	51.11	69.48

References

1. Banks PA, Bollen TL, Dervenis C, et al. Classification of acute pancreatitis--2012: revision of the Atlanta classification and definitions by international consensus. Gut 2013;62(1):102-111.
2. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29(1):15-21.
3. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014;30(7):923-930.
4. Tarazona S, Furió-Tarí P, Turrà D, et al. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. Nucleic Acids Res 2015;43(21):e140.
5. Ringnér M. What is principal component analysis? Nat Biotechnol 2008;26(3):303-304.
6. Lough G, Kyriazakis I, Bergmann S, et al. Health trajectories reveal the dynamic contributions of host genetic resistance and tolerance to infection outcome. Proc Biol Sci 2015;282(1819).
7. Strauss T, von Maltitz MJ. Generalising Ward's Method for Use with Manhattan Distances. PLoS One 2017;12(1):e0168288.
8. Miyamoto S, Abe R, Endo Y, et al. Ward method of hierarchical clustering for non-Euclidean similarity measures. In: 7th International Conference of Soft Computing and Pattern Recognition (SoCPaR); 2015; 2015.
9. Szekely GJ, Rizzo ML. Hierarchical Clustering via Joint Between-Within Distances: Extending Ward's Minimum Variance Method. Journal of Classification 2005;22(2):151-183.
10. Kanehisa M, Furumichi M, Tanabe M, et al. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 2017;45(D1):D353-D361.
11. Luo W, Friedman MS, Shedden K, et al. GAGE: generally applicable gene set enrichment for pathway analysis. BMC Bioinformatics 2009;10:161.
12. Xia J, Wishart DS. Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. Curr Protoc Bioinformatics 2016;55:14.10.11-14.10.91.
13. Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. Nature 2014;507(7493):455-461.
14. Lizio M, Harshbarger J, Shimoji H, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol 2015;16:22.
15. Il-Gyo C, Chi-Hyuck J. Performance of some variable selection methods when multicollinearity is present. ChemometrIntell Lab 2005;78:103-112.
16. Sweeney TE, Azad TD, Donato M, et al. Unsupervised Analysis of Transcriptomics in Bacterial Sepsis Across Multiple Datasets Reveals Three Robust Clusters. Crit Care Med 2018;46(6):915-925.
17. Calfee CS, Delucchi K, Parsons PE, et al. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. Lancet Respir Med 2014;2(8):611-620.
18. Rouseeuw P. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 1987;20:53-65.
19. Hanada K. Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. Biochim Biophys Acta 2003;1632(1-3):16-30.
20. Goldstein DS, Swoboda KJ, Miles JM, et al. Sources and physiological significance of plasma dopamine sulfate. J Clin Endocrinol Metab 1999;84(7):2523-2531.
21. Jetter A, Kinzig-Schippers M, Illauer M, et al. Phenotyping of N-acetyltransferase type 2 by caffeine from uncontrolled dietary exposure. Eur J Clin Pharmacol 2004;60(1):17-21.

22. Lankisch TO, Behrens G, Ehmer U, et al. Gilbert's syndrome and hyperbilirubinemia in protease inhibitor therapy--an extended haplotype of genetic variants increases risk in indinavir treatment. J Hepatol 2009;50(5):1010-1018.
23. Lankisch TO, Gillman TC, Erichsen TJ, et al. Aryl hydrocarbon receptor-mediated regulation of the human estrogen and bile acid UDP-glucuronosyltransferase 1A3 gene. Arch Toxicol 2008;82(9):573-582.
24. Mondry A. Antiglucocorticoid RU38486 reduces net protein catabolism in experimental acute renal failure. BMC Nephrol 2005;6:2.
25. Hawke TJ, Atkinson DJ, Kanatous SB, et al. Xin, an actin binding protein, is expressed within muscle satellite cells and newly regenerated skeletal muscle fibers. Am J Physiol Cell Physiol 2007;293(5):C1636-1644.
26. Okazaki T, Higuchi M, Takeda K, et al. The ASK family kinases differentially mediate induction of type I interferon and apoptosis during the antiviral response. Sci Signal 2015;8(388):ra78.
27. Taniguchi Y, Komatsu N, Moriuchi T. Overexpression of the HOX4A (HOXD3) homeobox gene in human erythroleukemia HEL cells results in altered adhesive properties. Blood 1995;85(10):2786-2794.
28. Hirata Y, Katagiri K, Nagaoka K, et al. TRIM48 Promotes ASK1 Activation and Cell Death through Ubiquitination-Dependent Degradation of the ASK1-Negative Regulator PRMT1. Cell Rep 2017;21(9):2447-2457.
29. Darnaud M, Dos Santos A, Gonzalez P, et al. Enteric Delivery of Regenerating Family Member 3 alpha Alters the Intestinal Microbiota and Controls Inflammation in Mice With Colitis. Gastroenterology 2018;154(4):1009-1023.e1014.
30. Delibegovic M, Armstrong CG, Dobbie L, et al. Disruption of the striated muscle glycogen targeting subunit PPP1R3A of protein phosphatase 1 leads to increased weight gain, fat deposition, and development of insulin resistance. Diabetes 2003;52(3):596-604.